

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 15:54:52 ; Search time 97.5789 Seconds

Sequence: (without alignments) (221.222 Million cell updates/sec)

Title: US-10-019-219-1
Perfect score: 912
Sequence: TVVRLFLAWLPCNMVPCNLP.....WAACGARVKRRFLQLTLSR 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Post-processing: Minimum Match 0%
Maximum Match 100%

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Database : A_Genesed_101002:*

1: /SIBS2/gcdata/geneseq/geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1989.DAT:*

10: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1990.DAT:*

11: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1991.DAT:*

12: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1992.DAT:*

13: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1993.DAT:*

14: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1994.DAT:*

15: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1995.DAT:*

16: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1996.DAT:*

17: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1997.DAT:*

18: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1998.DAT:*

19: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA1999.DAT:*

20: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA2000.DAT:*

21: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA2001.DAT:*

22: /SIBS2/gcdata/geneseq/geneseqp-emb1/AA2002.DAT:*

11 96.5 10.6 387 19 AAW81589
12 96 10.2 408 23 ABG5993
13 95.5 10.5 264 22 ABG05131
14 93.5 10.3 180 22 ABG12827
15 93.5 10.1 180 22 ABG14843
16 93.5 10.3 1061 20 AAU87504
17 93.5 10.3 1312 20 AAU87503
18 93 10.2 145 21 AX87201
19 93 10.2 145 22 AX806178
20 93 10.2 145 23 ABG05146
21 92.5 10.1 119 22 AAU5360
22 92.5 10.1 331 19 AAU81593
23 91.5 10.0 138 22 AAU64233
24 91 10.0 175 22 AAU87255
25 90.5 9.9 243 22 AAU60496
26 90.5 9.9 440 22 ABG05146
27 89.5 9.8 60 13 ABR24223
28 89.5 9.8 167 22 ABG19252
29 89 9.8 247 22 AAU46904
30 89 9.8 265 22 AAU43550
31 88 9.7 280 21 ABG34201
32 88.5 9.7 346 21 ABG40817
33 88.5 9.7 637 20 ABY04993
34 88.5 9.7 1232 21 ABZ26239
35 88.5 9.7 1232 23 AAU47961
36 88 9.6 106 23 ABP01922
37 88 9.6 207 22 ABZ02687
38 88 9.6 233 22 ABM86095
39 88 9.6 1021 23 ABT9496
40 88 9.6 1170 23 AAU79500
41 88 9.6 1224 23 AAU015254
42 88 9.6 1224 23 ABT7497
43 88 9.6 1224 23 AAU80153
44 88 9.6 1224 23 AAU72891
45 88 9.6 1236 23 AAU9886

ALIGNMENTS

RESULT 1	ID	Description
	AAB31701	AAB31701 standard; peptide; 162 AA.
XX	AAB31701;	
XX	DT	30-APR-2001 (first entry)
XX	DE	Peptide fragment of a human intestinal carboxylesterase (ICE).
XX	KW	Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL; cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma; tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
XX	OS	Homo sapiens.
XX	PD	W0200100784-A2.
XX	PD	04-JAN-2001.
XX	PF	27-JUN-2000; 2000WO-FR01791.
XX	PR	28-JUN-1999; 99FR-0008224.
XX	PA	(INSR) INST ROUSSY GUSTAVE.
XX	PI	Ronsin C, Scott V, Triebel F;
XX	DR	WPI; 2001-11243/12.
XX	PT	New peptides and its encoding nucleic acid derived from intestinal carboxylesterase, useful as immunostimulants for treating cancer.
XX	PT	Propionibacterium

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	912	100.0	162 22 AAB31701	Peptide fragment of protein encoded by Propionibacterium Novel human dihydrofolate reductase
2	912	100.0	166 22 AAB31703	Novel human dihydrofolate reductase
3	103.5	11.3	243 22 AAU46489	Novel human dihydrofolate reductase
4	101.5	11.1	270 22 ARG30149	Novel human dihydrofolate reductase
5	11.1	10.0	306 22 ARG19522	Novel human dihydrofolate reductase
6	98	10.7	241 20 AAY73982	Human prostrate tumour
7	98	10.7	470 23 AAU91561	Amino acid sequence
8	97	10.6	157 22 ARU49370	Propionibacterium
9	97	10.6	332 21 AAY82703	Tick derived cyste
10	96.5	10.6	271 22 AAU58489	Propionibacterium

PS Claim 1; Page 3; 53p; French.

XX The present sequence is derived from a human intestinal carboxylesterase
 CC polypeptide. ICE induces specific-specific cytotoxic T lymphocytes
 CC (ICE) polypeptide. ICE induces specific-specific cytotoxic T lymphocytes
 CC (CTL) and secretion by these CTL of cytotoxic factors, e.g.
 CC interleukin-2, interferon gamma and tumour necrosis factor. ICE
 CC polypeptides and polynucleotides are used for treating cancer, by in
 vivo or in vitro immunisation, particularly solid cancers and most
 especially hepatocarcinoma or adenocarcinoma of colon and kidney. They
 CC are also used to stimulate the immune system, and to increase, in
 culture, the production of associated-associated CTL, for reinjection,
 CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic
 CC cells loaded with ICE are used to stimulate the immune system, and to increase, in
 CC culture, the production of associated-associated CTL, for reinjection,
 CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic
 CC cells loaded with ICE are used to induce such CTL in cultures.

SQ Sequence 162 AA;

Query Match	Score	DB	Length
Best Local Similarity	100.0%	22	162
Matches	162;		
AC	Conservative	0;	Mismatches
XX	Indels	0;	Gaps

Qy 1 TVVRIFLAWLPCMMVPCWLPWRWTWWSSSSSTAWNSWASSALESTSTOPATGATWKLYHA 60
 Db 1 TVVRIFLAWLPCMMVPCWLPWRWTWWSSSSSTAWNSWASSALESTSTOPATGATWKLYHA 60

Qy 61 GSSRISPTIATELTIVSPFLSLRVARVCRLCPYKPSOSTERSWRYAWPSCASLPQ 120
 Db 61 GSSRISPTIATELTIVSPFLSLRVARVCRLCPYKPSOSTERSWRYAWPSCASLPQ 120

Qy 121 LMSSPRWWPTCLPVTKLTPWRWWAACGARYRKRRFLQLTSLSR 162
 Db 121 LMSSPRWWPTCLPVTKLTPWRWWAACGARYRKRRFLQLTSLSR 162

RESULT 2

AAB31703 AAB31703 standard; Protein: 166 AA.

AC AAB31703;
 XX 30-APR-2001 (first entry)

DE Protein encoded by an intestinal carboxylesterase (ICE) cDNA.
 XX Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;
 KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
 KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
 OS Homo sapiens.

XX WO200100784-A2.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-FR01791.

XX 28-JUN-1999; 99FR-0008224.
 PR (INSR) INST ROUSSEY GUSTAVE.

PA Ronsin C, Scott V, Triebel F;
 XX DR WPI; 2001-112443/12.
 DR N-PSDB; AAF25258.

PS New peptides and its encoding nucleic acid derived from intestinal
 PT carboxylesterase, useful as immunostimulants for treating cancer -
 Disclosure; fig 8A; 53p; French.

CC The present sequence is encoded by the coding region of human intestinal
 CC carboxylesterase (ICE) gene. ICE induces specific-specific cytotoxic T
 CC lymphocytes (CTL) and secretion by these CTL of cytotoxic factors,
 e.g. interleukin-2, interferon gamma and tumour necrosis factor. ICE

CC polypeptides and polynucleotides are used for treating cancer, by in
 vivo or in vitro immunisation, particularly solid cancers and most
 especially hepatocarcinoma or adenocarcinoma of colon and kidney. They

CC are also used to stimulate the immune system, and to increase, in
 culture, the production of associated-associated CTL, for reinjection,
 CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic
 CC cells loaded with ICE are used to stimulate the immune system, and to increase, in
 CC culture, the production of associated-associated CTL, for reinjection,
 CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic
 CC cells loaded with ICE are used to induce such CTL in cultures.

SQ Sequence 166 AA;

Query Match	Score	DB	Length
Best Local Similarity	100.0%	22	166
Matches	166;		
AC	Conservative	0;	Mismatches
XX	Indels	0;	Gaps

Qy 1 TVVRIFLAWLPCMMVPCWLPWRWTWWSSSSSTAWNSWASSALESTSTOPATGATWKLYHA 60
 Db 5 TVVRIFLAWLPCMMVPCWLPWRWTWWSSSSSTAWNSWASSALESTSTOPATGATWKLYHA 64

Qy 61 GSSRISPTIATELTIVSPFLSLRVARVCRLCPYKPSOSTERSWRYAWPSCASLPQ 120
 Db 65 GSSRISPTIATELTIVSPFLSLRVARVCRLCPYKPSOSTERSWRYAWPSCASLPQ 124

Qy 121 LMSSPRWWPTCLPVTKLTPWRWWAACGARYRKRRFLQLTSLSR 162
 Db 125 LMSSPRWWPTCLPVTKLTPWRWWAACGARYRKRRFLQLTSLSR 166

RESULT 3

AAU64689 AAU64689 standard; Protein: 243 AA.
 AC AAU64689;
 XX DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #7385.
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX PN WO20010181581-A2.

XX PD 01-NOV-2001.

XX PP 20-APR-2001; 2001WO-US12865.

XX PR 21-RPR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208041P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI) CORIXA CORP.

XX PR Skeley YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PA L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX PI WPI; 2001-616747/1.

XX DR DR N-PSDB; AAS5932.

XX PS Example 1; SEQ ID No 7684; 1069pp; English.

CC Sequences AAU39105-AAU60017 represent Propionibacterium acnes immunogenic
 CC polyPeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

SQ	Sequence	243 AA:
Query Match 11.3%; Score 103.5; DB 22; Length 243; Best Local Similarity 28.1%; Pred. No. 0.95%; Mismatches 62; Indels 45; Gaps 47; Conservative 13; MisMatches 62; Indels 45;		
QY	8 AW-LPCMV--PCWLP WRTWWSSSTAWSSALLETS-TOPATGATWKWLHYAG 61	
Db	6 AWSSTCHSWSVTRPGWMETRTMPRSSRRAWSARSYFDAYVPILAGV----- 55	
QY	62 SSRISPTPLEATLTVSPFLA-----SLRVARYAVCIRLLCPPYKPDGSSTERPSWRVAVPSCPAS 116	
Db	56 -RRNPDPPELIMTSPWLARSSGSTRRVRSATMLNTRRHSSR--GWEPISPGDVAL 111	
QY	117 LPQLQMSPPRW-WPTCLPV-----TKitLRPWW 143	
Db	112 TPAALLRASRWSSEAVAVASSSVRSITKGERALGSSDTTLSTASWW 158	

RESULT 4

ID	ABC30149
AC	ABC30149 standard; Protein; 270 AA.
XX	
DE	Novel human diagnostic protein #30140.
XX	
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
KW	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PP	30-MAR-2001; 2001WO-US08631.
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PT	Drmancic RT, Liu C, Tang YT;
XX	
DR	WPI- 2001-639362/73. N-PSSDB; AAS94336.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
XX	
PS	Claim 20; SEQ ID No 60508; 103pp; English.
XX	

XX
 PT diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits and to assess
 CC biodiversity -
 PS claim 20; SEQ ID No 49911; 103pp; English.
 The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosomes
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tag
 CC for identifying expressed genes. (I) is useful in gene therapy technique
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and
 CC a food supplement. (II) and its binding partners are useful in medical
 imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity
 CC in the polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AB00010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 Note: The sequence data for this patent did not appear in the printed
 specification, but was obtained in electronic format directly from WIPO
 at [ftp://www.wipo.int/pdb/published_pct_sequences](http://www.wipo.int/pdb/published_pct_sequences).
 XX
 SQ sequence 306 AA:
 Query Match 11.0%; Score 100; DB 22; Length 306;
 Best Local Similarity 25.0%; Pred. No. 0; 16'; Mismatches 57; Indels 42; Gaps
 Matches 40; Conservative 15; Mismatches 57; Indels 42; Gaps
 QY 10 LPCNMVPCMLP-----WRTWWN-----SSSTAWYWASSALETSQPATGATW 53
 Db 111 LPGQMRSRHHMSPKIRPANTWRTSWNALPQAPGLSPGAQSWSFVPRV-TPTPGTCTCPT 168
 QY 54 TKWLHYAGSSRISPTLEATIVSPFLASLR-----VARVCLRLICPPYPKDSSTE 103
 Db 169 PSWNSWRSSRTSSL-----NPFTESLKNLSRSTSKEFGWSQSAMATPCPULLPPASRVE 222
 QY 104 PSWNRVAWPSCPASLPAQQLMSSPRWNWPTCLPWTKL 137
 Db 223 -LWPETWMP-----LGQGHATSSOW--KCVPERKL 248
 RESULT 6
 AAV73982
 ID AAV73982 standard; Protein; 241 AA.
 XX
 AC AAV73982;
 XX
 DT 14-MAR-2000 (first entry)
 DE Human prostate tumor EST fragment derived protein #169.
 XX
 KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
 KW treatment.
 XX
 OS Homo sapiens.
 XX
 PN DE19820190-A1.
 XX
 PR 04-NOV-1999.
 XX
 PF 28-APR-1998; 98DE-1020190.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX DR WPI; 1999-521385/54.
 XX N/P SDB; AA252913;
 XX
 PT New human nucleic acid sequences from pancreatic tumors, and related
 PT proteins
 XX
 PS Claim 23; Page 380; 502pp; German.
 XX
 CC This invention describes novel polypeptides and their encoding nucleic
 CC acids derived from human pancreatic tumor tissue which have cytostatic
 CC activity. The sequences are also useful in producing pharmaceutical
 CC compositions for treatment of pancreatic tumors. AA73814-Y7452
 CC represent protein fragments encoded by the human pancreatic tumor cDNA
 CC library derived expressed sequence tag (EST) sequences represented in
 XX AA252858-Z53014.
 XX
 SQ Sequence 241 AA;
 Query Match 10.7%; Score 98; DB 20; Length 241;
 Best Local Similarity 25.7%; Pred. No. 0.18;
 Matches 45; Conservative 10; Mismatches 64; Indels 56; Gaps 9;
 QY 9 WIPLPCM-----VPCWLWPTWWMSSSTAWNSHASSALETSTHOPATGATWTWKLH 58
 Db 38 WIPLAALRNKVEAPERWSPWCW-AWQWQ----WEPW---LWGWPPEPGIGRTSFERQ 87
 QY 59 YA-----GSSRISPLEATLWVSPELASLRVARVCRLICPPY-----PKDSTE 103
 Db 88 EATGQTLLACQTSRTPGNLPELPTWEPRLSRRLHPSEEKKSLLPPRLPQRQPNKRQGH 147
 QY 104 PSRVRVAD--SCPASLPAQLMSPRWPMTCPLVTKTLRPPWAAACGARKRKFL 155
 Db 148 PRKPRWPKTSC-----SSPAPWPP-----RPTAPRKPRCRRL 184
 RESULT 7
 AAU91561
 ID AAU91561 standard; protein: 470 AA..
 XX
 AC AAU91561;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Amino acid sequence of human PHOR1-F5D6 splice variant C (frame 2).
 XX Human; PHOR1-F5D6; prostate cancer; cytostatic.
 OS Homo sapiens.
 XX
 PN WO200214501-A2.
 XX
 PR 21-FEB-2002.
 XX
 PF 17-AUG-2001; 2001WO-US25862.
 XX
 PD 17-AUG-2000; 2000US-226241P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
 XX Jakobovits A;
 XX DR; 2002-269193/31.
 XX
 PT Monitoring PHOR1-A11/PHOR1-F5D6 gene products for monitoring presence
 PT of cancer in subject, by determining status of PHOR1-A11/PHOR1-F5D6
 PT gene products in tissue sample from subject and comparing it to normal
 PT sample -
 XX Example 45; Page 217; 250pp; English.
 XX
 CC The present invention relates to the isolation of novel human genes

CC designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The CC gene encoding PHOR1-All map to chromosome 1q23, and the gene encoding CC PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6 CC polypeptide and polypeptide sequences are useful in diagnostic and therapeutic methods, and compositions for various cancers such as CC prostate cancer. The sequences are useful for inhibiting the growth of CC cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating CC P. acnes. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof CC can be used to elicit an immune response. The present sequence represents an amino acid sequence from the translation of the DNA CC sequence for human PHOR1-F5D6 splice variant C.

SEQ Sequence 470 AA;

Query Match Similarity 10.7%; Score 98; DB 23; Length 470;
Best Local Similarity 25.1%; Pred. No. 0; 4; Matches 42; Conservative 15; Mismatches 46; Indels 64; Gaps 12;

Qy 6 FLAWLPCMMVCPWLPWTTWWSSPAW-VEWMS---SALETSSTOPATGATW--TKWL 57
Db 246 FVCFVCFLVVKWKX1RQ - WSQSSYWDFSWAQGPRCSSLGSPSMSSPCWGMCPSW- 302

Qy 58 HYAGGSRISPLPLEATLTVSPFLASLRVAVRLRLCPPYPPDSSTPSWRVAPWPSPL 117
Db 303 ---GSSHWT--DST-----PPCTSSHT-----WPSSTSPM 329

Qy 118 PAQLMSSPRRWMP-CL----PVTKLILRPM-----WAAC 146
Db 330 PA-TOPCRWXTSCTIQSPSPSPLAAXHRPFSFXVLUHILNASCWCXC 374

RESULT 8
AAU49370 ID AAU49370 standard; Protein: 157 AA.

XX AC AAU49370;
XX DT 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #10266.
XX DE

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX OS W0200181581-A2.

PN 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR WPI; 2001-616774/71.
DR N-PSDB; AAS59545.

XX Propionibacterium acnes polypeptides and nucleic acids useful for PT vaccinating against and diagnosing infections, especially useful for PT treating acne vulgaris -
Example 1; SEQ ID NO 10565; 1059pp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in CC the treatment, prevention and diagnosis of medical conditions caused by CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis. CC P. acnes is also involved in infections of bone, joints and the central CC nervous system, however it is particularly involved in the inflammatory CC lesions associated with acne vulgaris. A method for detecting the CC presence or absence of P. acnes in a patient comprises contacting a CC sample with a binding agent that binds to the proteins of the invention CC and determining the amount of bound protein in the sample. The CC polypeptides may be used as antigens in the production of antibodies CC specific for P. acnes proteins. These antibodies can be used to CC downregulate expression and activity of P. acnes polypeptides and CC therefore treat P. acnes infections. The antibodies may also be used as CC diagnostic agents for determining P. acnes presence, for example, by CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp://wipo.int/pub/published_pct_sequences.

SEQ Sequence 157 AA;

Query Match Similarity 10.6%; Score 97; DB 22; Length 157;
Best Local Similarity 28.9%; Pred. No. 0; 14; Matches 39; Conservative 17; Mismatches 43; Indels 36; Gaps 9;

Qy 26 WSSSTTAW--VWASSALETSSTOPATGATWTKWLHYAGGSSRTSP-----TEATLTVPFL 79
Db 1 WRICAPAWNSAVRWARATCA-PSTASIWSR-----SRMPPARTRLESPTVFPSP 52

Qy 80 ASLRVARVCLRICPPYPKDSTE--PSW - RVAMPSCPASLPAQLMSSPRKWPCLPV 134
Db 53 TSMRSSRL-VSTIVSSLPRRSSTRRSVSSWLRRGSTP-----SSRSRWPTMLRL 100

Qy 135 TKUTLRP---WVA 144
Db 101 RLWLRPLRARVWA 115

RESULT 9

AY82703 ID AY82703 standard; Protein: 332 AA.
XX AC AY82703;

XX DT 10-AUG-2000 (first entry)

XX DE Tick derived cysteine protease protein sequence #2.
XX KW Tick; vaccine; infection; salivary gland antigen; immunogen;
KW serine protease; cysteine protease; blood sucker.
XX OS Haemaphysalis longicornis.
XX PN JP2000083677-A..

XX PD 28-MAR-2000.
XX PF 17-SEP-1998; 98JP-0281932.

XX PR 17-SEP-1998; 98JP-0281932.
XX PA (FARB) BAYER KK.

XX DR WPI; 2000-296340/26.

DR N-PSDB; AAS59635.

XX A gene encoding tick salivary gland antigen - useful as a vaccine for PT protecting animals from tick-carried infections
XX
PS Claim 11; Page 17; 29pp; Japanese.

CC The present sequence represents a tick derived cysteine protease. The

CC present invention also describes a tick salivary gland antigen related
 CC immunogen and a tick derived serine protease. A nucleotide sequence
 CC encoding any of the above proteins can be used in a vaccine against
 CC tick carried infections for domestic animals such as cattle.
 XX

SQ Sequence 332 AA;

Query Match Similarity 10.6%; Score 97; DB 21; Length 332;
 Best Local Similarity 29.5%; Pred. No. 0.33; Matches 36; Conservative 36; Mismatches 10; Indels 42; Gaps 6;

QY 17 CWLFW-RWWHSSSTAWNSWASSALESTQPATGAWTKWLHYAGSSRISPLEAELT 74
 158 CWWHSSSTAWNSWASSALESTQPATGAWTKWLHYAGSSRISPLEAELT 74
 Db 17 CWLFW-RWWHSSSTAWNSWASSALESTQPATGAWTKWLHYAGSSRISPLEAELT 74
 158 CWWHSSSTAWNSWASSALESTQPATGAWTKWLHYAGSSRISPLEAELT 74

QY 75 VSPFLASLRVARVCLRLCPPY-PKDSSTEPSWVAVWPSCPASLPAQLMSSPRWPTCLP 133
 206 TA---RTRATSR--PPWGPGTAGLTSPETRRSC----RRWPTVGP 246
 Db 134 VT 135
 247 VS 248

RESULT 10

AAU58489 ID AAU58489 standard; Protein; 271 AA.

AC ACU58489;

XX DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #19385.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; BLIST; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

XX OS WO200181581-A2.

XX PD 01-NOV-2001.

XX PR 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORT-) CORIXA CORP.

PI Skeky YMW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616747/11.

DR N-FSDB; AAU55591.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for vaccinating against and diagnosing infections, especially useful for treating acne vulgaris - Example 1; SEQ ID No 19684; 1069pp; English.

CC Sequences AAU39105-AAU88017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteonecrosis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention CC and determining the amount of bound protein in the sample. The CC polypeptides may be used as antigens in the production of antibodies CC specific for P. acnes proteins. These antibodies can be used to CC downregulate expression and activity of P. acnes polypeptides and CC therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at [ftp.wipo.int/pdb/published/pct/sequences](http://wipo.int/pdb/published/pct/sequences).

SQ Sequence 271 AA;

Query Match Similarity 10.6%; Score 96.5; DB 22; Length 271;
 Best Local Similarity 24.2%; Pred. No. 0.29; Matches 40; Conservative 40; Mismatches 58; Indels 43; Gaps 9;

QY 17 CWLFWRTWWS---SSTAWNSWASSALESTQPATGAWTKWLHYAGSSRISPLEAELT 74
 17 CWLFWRTWWS---SSTAWNSWASSALESTQPATGAWTKWLHYAGSSRISPLEAELT 74
 Db 67 SRPSICSTRSSVAPRTRPAP-----AYRNCLISTRFRILAPSSICSCVW 115

QY 128 WPTCL-----PVTKITLRPWAAACGARVVRKFLLTLSR 162
 Db 116 LPWSVKRSSTDLNKNSASSITKOMWLARTASANS--SLSPLR 157

RESULT 11

AAW81589 ID AAW81589 standard; Protein; 387 AA.

XX AC AAW81589;

XX DT 09-FEB-1999. (first entry)

DE Protein encoded by human UCP3 gene reading frame 2.

XX KW Uncoupling protein 3; UCP3; thermogenesis; mammal; enhancer; drug; protein catabolism; anti-obesity; inhibitor; muscle wasting; infection; HIV; cancer; tumour cachexia; muscle disease; muscular dystrophy; non-insulin dependent diabetes mellitus; diagnosis; human.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT FT 1. 387

FT FT /note= "the encoding reading frame has internal stop codons which are not indicated in this protein"

PN W09845438-A1.

XX PN

PD 15-OCT-1998.

XX PF 08-APR-1998; 98WO-US06959.

XX PR 15-JUL-1997; 97US-0892745.

PR 09-APR-1997; 97US-0043447.

PR 12-MAY-1997; 97US-0046254.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PI Flier JS, Lowell BB;

XX WPI; 1998-594483/50.

DR N-FSDB; AAU71710.

PT New isolated uncoupling protein, UCP-3 - used to develop products

Db 92 WC--ACTGPRATSPRAESTPITSQSCPASTASP-----TPCSTASSARRPPTGTWMPAS 143
 Qy 109 ---AWPSCPASLPAQLMSSPRWMP 129
 Db 144 AGPAWPSSPATGPAGPGRPTKWP 167

RESULT 15

Db 92 WC--ACTGPRATSPRAESTPITSQSCPASTASP-----TPCSTASSARRPPTGTWMPAS 143
 ID ABG14843 standard; Protein; 180 AA.
 XX
 AC ABG14843;
 XX
 DR 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #14834.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.

XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PT Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/773.
 XX
 PT N-PSDB; AAS79030.
 PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
 XX
 PS Claim 20; SEQ ID NO 45202; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG3037 represent novel human diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 XX
 Sequence 180 AA;

Query Match 10.3%; Score 93.5; DB 22; Length 180;
 Best Local Similarity 27.1%; Pred. No. 0..35; Indels 49; Gaps 8;
 Matches 39; Conservative 8; Mismatches 48;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:08:52 ; Search time 35.0526 Seconds

Sequence: {without alignments} 135.982 Million cell updates/sec

Title: US-10-019-219-1

Perfect score: 912

Sequence: 1 TTVRLFLAWLPCMMVPCMLP.....WAACGARYKRRFQLTLSR 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB_pep:*

2: /cgn2_6/ptodata/1/iaa/5B_COMB_pep:*

3: /cgn2_6/ptodata/1/iaa/6A_COMBO_pep:*

4: /cgn2_6/ptodata/1/iaa/6B_COMB_pep:*

5: /cgn2_6/ptodata/1/iaa/PTCNS_COMB_pep:*

6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID

Description

Sequence 4, Appl1

Sequence 81, Appl1

Sequence 16, Appl1

Sequence 12, Appl1

Sequence 2, Appl1

RESULT 1

US-09-930-872-4

Sequence 4, Application US/09930872

Patent No. 6443888

GENERAL INFORMATION:

APPLICANT: Fiddle, Carl Johan

APPLICANT: Hibun, Erin

TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the FILE REFERENCE: LEX-0219-A

CURRENT APPLICATION NUMBER: US/09/930,872

CURRENT FILING DATE: 2001-08-14

PRIOR APPLICATION NUMBER: US 60/225,852

PRIOR FILING DATE: 2000-08-16

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 4

LENGTH: 1224

TYPE: PRT

ORGANISM: homo sapiens

US-09-930-872-4

Query Match 9.6%; Score 88; DB 4; Length 1224;

Best Local Similarity 25.5%; Pred. No. 1,3; Mismatches 18; Indels 34; Gaps 7;

Matches 37; Conservative 18; Mismatches 56; Indels 34; Gaps 7;

QY 12 CMAVPCWLPWRWWWSSTANVSWASSALESTSTOPATGATWTWKLYAGSSRSRISPLEA 71

Db 1042 CLIQORCHKKPLQLWLS--AN-SOCSYTCERGTOKFLKCAEKKY--SGKYRELASKKC 1095

QY 72 TLTVPSPFLASLRVAVRLRLCPPYKDSTOPSWRYAWPSCPASLPAQLMSSPRWMPTC 131

Db 1096 SHLPKP--SLELERACAPLCPPRHPPFAAGPS-RGSWFASP-----WSQC 1138

QY 132 LPVTKLTLRPPWNACGARYKRAFLQ 156

Db 1139 T ----- ASCGGGVQTSVQ 1152

RESULT 2

US-08-469-260A-81

Sequence 3, Appl1

; Sequence 81, Application US/08469260A.

Patent No. 6451578

GENERAL INFORMATION:

APPLICANT: JOHN N. SIMONS

APPLICANT: TAM J. PILOT-MATIAS

APPLICANT: GEORGE J. DAWSON

APPLICANT: GEORGE G. SCHLAUDER

APPLICANT: SURESH M. DESAI

APPLICANT: THOMAS P. LEARY

APPLICANT: ANTHONY SCOTT MUERHOF

APPLICANT: JAMES C. ERKER
 APPLICANT: SHERI L. BUIJK
 APPLICANT: ISRA MUSHHARAR
 TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
 NUMBER OF SEQUENCES: 716
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
 STREET: 100 ABBOTT PARK ROAD
 CITY: ABBOTT PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/469,260A
 FILING DATE: 17-MAR-1997
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/08/424,550
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: POREMBSKI, PRISCILLA E.
 REGISTRATION NUMBER: 33,207
 REFERENCE/DOCKET NUMBER: 5527.PC.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-3665
 TELEFAX: 708-938-2623
 INFORMATION FOR SEQ ID NO: 81:
 LENGTH: 449 amino acids
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1-22 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ; US-08-469-260A-81
 Query Match 9.6%; score 87.5; DB 4; Length 1422;
 Best Local Similarity 21.1%; Pred. No. 1.7; Mismatches 55; Indels 95; Gaps 9;
 Matches 45; Conservative 18; Mismatches 55; Indels 95; Gaps 9;
 Qy 10 LPCMNVPCULPWRWWSSSTAWVWASSALETSTQPATGATWKWVHYAGSSRI-SPT 68
 Db 594 LPVILSPQWLTKRSWRSHHS-----FPPRPWMLQLQTSKRVQSPQ 634
 Qy 69 LEAT-----LTVSPFL-----ASLRVAR-----VCLRLHCPYY--- 96
 Db 635 LVLSHWKPPLKNUIPFLGIMQLOSSLSXSIAVAXSLYLTIPHLHACLLSLRVILPHYLTR 694
 Qy 97 -----PKDSSTEPSHRVAMP-----SCAS----- 116
 Db 695 SKCSCHYLEAQRLPSQTLTDXHWRXWPGILXEQOLVHGRWVLSTLCXAMLAHPHLLAX 754
 Qy 117 -LPQLMSSPRWPTCLPVTKLTPR- WWAAC 146
 Db 755 HLNKXWSGXLWSSLVAVSTPRSRQELWASC 787
 RESULT 3
 US 08-819-458A-16
 Sequence 16, Application US/08819458A
 ; Patent No. 5891669
 GENERAL INFORMATION:
 APPLICANT: Jensen, Einer B.
 APPLICANT: Cherry, Joel
 APPLICANT: Eliod, Susan L.
 TITLE OF INVENTION: Methods For Producing Polypeptides
 NUMBER OF SEQUENCES: 20
 ;
 RESULT 4
 US-08-660-963-12
 Sequence 12, Application US/08660963
 ; Patent No. 5852187
 GENERAL INFORMATION:
 APPLICANT: Thorner, Michael O.
 APPLICANT: Gaylann, Bruce D.
 APPLICANT: Horikawa, Reiko
 APPLICANT: Lyons Jr., Charles E.
 TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: POPHAM, RALK, SCHNOORICH & KAUFMAN, LTD.
 STREET: Metropolitan Square Building, Suite 800, 1450
 G. Street
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/660,963
 FILING DATE: 12-JUN-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: O'Shaughnessy, Brian P.
 REGISTRATION NUMBER: 32,747
 REFERENCE/DOCKET NUMBER: 18046-036
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-824-8000
 TELEFAX: 202-824-8199
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 498 amino acids
 TYPE: amino acid
 STRAND EDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-660-963-12

Query Match 9.3%; Score 84.5; DB 2; Length 498;
 Best Local Similarity 26.6%; Pred. No. 0.93; Matches 41; Conservative 14; Mismatches 40; Indels 59; Gaps 9;

Qy 20 PWRTW----WWSSTAWVWSASSALENTSTOPA-----"NGATWT 54
 Db 179 PWATASRLQFWPPSSS--WSLUSGGSTAPGTISTPSCSPPLSSRQLCSRTPPSGTGRWT 236

Qy 55 KWLHYAGSSRSIPTPLETLTVSPFLASLRVARYCLVCLICPPYPKQSSTERSW-RYAWPSC 113
 Db 237 ---"TAASPLS-----CARLUPPLISRPTTSAGCWQKLCPTAS 270

Qy 114 PASLPAOLMSSPRWWPTCLPVTKLURPWWAG 147
 Db 271 PPHPRAQGGSSGGWF--SLPGGFLSSP--ACG 299

RESULT 5
 US-08-735-041A-2
 Sequence 2, Application US/08735041A
 ; Sequence 2, Application US/08735041A
 ; Patent No. 5914251
 ; GENERAL INFORMATION:
 ; APPLICANT: Farrell, Catherine L.
 ; APPLICANT: Martin, Francis H.
 ; APPLICANT: Yakowitz, Rachel
 ; TITLE OF INVENTION: PLACENTAL-DERIVED PROSTRATE GROWTH
 ; TITLE OF INVENTION: FACTOR
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Agen Inc.
 ; STREET: 1840 De Havilland Drive
 ; CITY: Thousand Oaks
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 91330-1789

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/190,476B
 FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/735,041
 ; FILING DATE: 22-OCT-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mazza, Richard J.
 ; REGISTRATION NUMBER: 27,657
 ; REFERENCE/DOCKET NUMBER: A-414
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 805.447.4112
 ; TELEFAX: 805.499.6751
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 539 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-190-476B-2

Query Match 8.8%; Score 80; DB 3; Length 539;
 Best Local Similarity 28.8%; Pred. No. 3; Matches 17; Conservative 9; Mismatches 11; Indels 22; Gaps 3;

Qy 17 CWLP-----WRTWWNSSSSTAWVWSASSALENTSTOPATGWT--KWLHYAG 61

RESULT 7
 US-09-190-889A-2
 Sequence 2, Application US/09190889A
 Patent No. 6075008
 GENERAL INFORMATION:
 APPLICANT: Farrell, Catherine L.
 APPLICANT: Martin, Francis H.
 APPLICANT: Yabkowitz, Rachel
 TITLE OF INVENTION: PLACENTAL-DERIVED PROSTRATE GROWTH FACTOR
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Augen Inc.
 STREET: 1840 De Havilland Drive
 CITY: Thousand Oaks
 STATE: California
 COUNTRY: USA
 ZIP: 91320-1789
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/190,938B
 FILING DATE: 12-No. 6197339-1598
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Mazza, Richard J.
 REGISTRATION NUMBER: 27,657
 REFERENCE/DOCKET NUMBER: A-414
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 805.447.4112
 TELEFAX: 805.499.6751
 INFORMATION FOR SEQ ID NO: 2:
 APPLICATION NUMBER: US 08/735,041
 FILING DATE: 22-OCT-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Mazza, Richard J.
 REGISTRATION NUMBER: 27,657
 REFERENCE/DOCKET NUMBER: A-414
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 805.447.4112
 FAX: 805.499.6751
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 539 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 REFERENCE/DOCKET NUMBER: A-414
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 805.447.4112
 FAX: 805.499.6751
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 539 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 STRANDEDNESS: single
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-190-938B-2
 Query Match 8.8%; Score 80; DB 4; Length 539;
 Best Local Similarity 28.8%; Pred. No. 3; Mismatches 11; Indels 22; Gaps 3;
 Matches 17; Conservative 9; MisMatches 11; Indels 22; Gaps 3;
 Oy 17 CWLPL-----WRTWWWWSSSTAWNSWASSALETSQPATGATWT--KWHYAG 61
 Db 398 CWLSEGLRRTASGALWRSAGAWAAESSW-----STSIRPAGGPSWIGRGQWIVQLG 448
RESULT 9
 PCT-US95-09261-2
 Sequence 2, Application PC/TUS9509261
 GENERAL INFORMATION:
 APPLICANT: NAME: BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM
 APPLICANT: STREET: 201 West 7th Street
 APPLICANT: CITY: Austin
 APPLICANT: STATE: Texas
 APPLICANT: COUNTRY: United States of America
 APPLICANT: POSTAL CODE: 78701
 APPLICANT: TELEPHONE NO.: (512)499-4462
 APPLICANT: TELEFAX: (512)499-4523
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE EXPRESSION OF
 TITLE OF INVENTION: A BONE AND PROSTATE DERIVED GROWTH FACTOR
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/09261
 FILING DATE: CONCURRENTLY HEREWITH
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/283,701

MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08480-774A-2

Query Match 8.7%; Score 79.5; DB 2; Length 142;
 Best Local Similarity 34.0%; Pred. No. 0.62; Mismatches 15; Indels 7; Gaps 2;
 Matches 17; Conservative 11; Misnmatches 15; Indels 7; Gaps 2;

RESULT 12
 US-08-545-809A-114
 Sequence 114, Application US/08545809A
 ; Patent No. 6036878

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku
 TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
 NUMBER OF SEQUENCES: 145
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson, P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/545,809A
 FILING DATE: 27-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP93/00603
 FILING DATE: 10-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.
 REGISTRATION NUMBER: 29,066
 REFERENCE/DOCKET NUMBER: 00501/004001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070
 TELEFAX: 617-542-8906
 TELEX: 200154

INFORMATION FOR SEQ ID NO: 114:

SEQUENCE CHARACTERISTICS:
 LENGTH: 117 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-545-809A-114

Query Match 8.7%; Score 79; DB 3; Length 117;
 Best Local Similarity 29.4%; Pred. No. 0.55; Mismatches 13; Indels 6; Gaps 1;

RESULT 13
 US-09-260-527-3

Sequence 3, Application US/09260527A
 ; Patent No. 6228599
 ; GENERAL INFORMATION:
 ; APPLICANT: Fish & Richardson, P.C.
 ; APPLICANT: Mikkelson, J.D.

APPLICANT: Willarts, W. G.
 TITLE OF INVENTION: ANTIBODY
 FILE REFERENCE: IYQ019_001AU5

CURRENT APPLICATION NUMBER: US/09/260,527A
 CURRENT FILING DATE: 1999-02-26
 NUMBER OF SEQ ID NOS: 7

SEQ ID NO 3
 LENGTH: 278
 TYPE: PRT
 ORGANISM: UNKNOWN

FEATURE:

OTHER INFORMATION: Anti-homogalacturonan specific antibodies selected from a naive phage display library known as the Synthetic scfv Library (#1) from the Centre for OTHER INFORMATION: Protein Engineering, MRC Centre, Cambridge, UK

RESULT 14
 US-08-851-362D-22
 ; Sequence 22, Application US/08851362D
 ; Patent No. 6225883

GENERAL INFORMATION:

APPLICANT: Jakobovits, Aya
 APPLICANT: Yang, Xiao-Dong
 APPLICANT: Gallo, Michael
 APPLICANT: Jia, Xiao-Chi

TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal Growth Factor Receptor

FILE REFERENCE: Cell 4.20

CURRENT APPLICATION NUMBER: US/08/851,362D

CURRENT FILING DATE: 1997-05-05

NUMBER OF SEQ ID NOS: 49

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 22
 LENGTH: 76
 TYPE: PRT
 ORGANISM: human

US-08-851-362D-22

Query Match 8.7%; Score 79; DB 3; Length 117;
 Best Local Similarity 29.4%; Pred. No. 0.55; Mismatches 13; Indels 6; Gaps 1;

RESULT 15
 US-08-545-809A-142
 ; Sequence 142, Application US/08545809A
 ; Patent No. 6036878

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku
 TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME

NUMBER OF SEQUENCES: 145

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson, P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA

APPLICANT: Mikkelson, J.D.

APPLICANT: Willarts, W. G.
 TITLE OF INVENTION: ANTIBODY
 FILE REFERENCE: IYQ019_001AU5

CURRENT APPLICATION NUMBER: US/09/260,527A
 CURRENT FILING DATE: 1999-02-26
 NUMBER OF SEQ ID NOS: 7

SEQ ID NO 3
 LENGTH: 278
 TYPE: PRT
 ORGANISM: UNKNOWN

FEATURE:

OTHER INFORMATION: Anti-homogalacturonan specific antibodies selected from a naive phage display library known as the Synthetic scfv Library (#1) from the Centre for OTHER INFORMATION: Protein Engineering, MRC Centre, Cambridge, UK

RESULT 16
 US-09-260-527-3

Query Match 8.7%; Score 79; DB 4; Length 278;
 Best Local Similarity 29.4%; Pred. No. 1.6; Mismatches 13; Indels 6; Gaps 1;

Matches 15; Conservative 11; Misnmatches 16; Indels 7; Gaps 2;

RESULT 17
 US-08-851-362D-22
 ; Sequence 22, Application US/08851362D
 ; Patent No. 6225883

GENERAL INFORMATION:

APPLICANT: Jakobovits, Aya
 APPLICANT: Yang, Xiao-Dong
 APPLICANT: Gallo, Michael
 APPLICANT: Jia, Xiao-Chi

TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal Growth Factor Receptor

FILE REFERENCE: Cell 4.20

CURRENT APPLICATION NUMBER: US/08/851,362D

CURRENT FILING DATE: 1997-05-05

NUMBER OF SEQ ID NOS: 49

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 22
 LENGTH: 76
 TYPE: PRT
 ORGANISM: human

US-08-851-362D-22

Query Match 8.7%; Score 79; DB 4; Length 76;
 Best Local Similarity 32.0%; Pred. No. 0.57; Mismatches 16; Indels 7; Gaps 2;

Matches 16; Conservative 11; Misnmatches 16; Indels 7; Gaps 2;

RESULT 18
 US-08-851-362D-22
 ; Sequence 22, Application US/08851362D
 ; Patent No. 6225883

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku
 TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME

NUMBER OF SEQUENCES: 145

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson, P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA

APPLICANT: Mikkelson, J.D.

COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29 066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 200154
TELEX:
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-545-809A-142

Query Match 8.4%; Score 76.5; DB 3; Length 118;
Best Local Similarity 32.0%; Pred. No. 1;
Matches 16; Conservative 11; Mismatches 16; Indels 7; Gaps 2;
QY 27 SSSSTAWWSSWASALERSTSTOPATGATWKWLWAGSSRSRISPTLEATLVS 76
DB 49 SSSYYW-SWIRO-----PPGKGLEWGYIYSGSTNYPNSIKSRVITIS 91

Search completed: March 27, 2003, 16:13:22
Job time : 36.0526 secs

(THIS PAGE BLANK (USPTO))

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:12:13 ; Search time 38.8421 Seconds

(without alignments) 244.921 Million cell updates/sec

Title: US-10-019-219-1
perfect score: 912
Sequence: 1 TVVRLFLAWLPCMVWPCWLDP.....WAACGAVRKRRFLOLTSR 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%, Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pupaa/us08_NEW_PUB.pep: *
2: /cgn2_6/ptodata/2/pupaa/ptc_NEW_PUB.pep: *
3: /cgn2_6/ptodata/2/pupaa/us06_NEW_PUB.pep: *
4: /cgn2_6/ptodata/2/pupaa/us07_NEW_PUB.pep: *
5: /cgn2_6/ptodata/2/pupaa/us07_PUBCOMB.pep: *
6: /cgn2_6/ptodata/2/pupaa/us07_PUBCOMB.pep: *
7: /cgn2_6/ptodata/2/pupaa/ptcNEW_PUBCOMB.pep: *
8: /cgn2_6/ptodata/2/pupaa/us08_PUBCOMB.pep: *
9: /cgn2_6/ptodata/2/pupaa/us09_NEW_PUB.pep: *
10: /cgn2_6/ptodata/2/pupaa/us09_PUBCOMB.pep: *
11: /cgn2_6/ptodata/2/pupaa/us10_NEW_PUB.pep: *
12: /cgn2_6/ptodata/2/pupaa/us10_PUBCOMB.pep: *
13: /cgn2_6/ptodata/2/pupaa/us60_NEW_PUB.pep: *
14: /cgn2_6/ptodata/2/pupaa/us60_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	9.3	10.2	145	US-09-984-271-240
2	9.1	10.0	265	Sequence 240, APP
3	8.8	9.6	1224	Sequence 77, APP
4	87.5	9.6	890	Sequence 4, APP
5	87.5	9.6	1422	Sequence 8, APP
6	85.5	9.4	890	Sequence 81, APP
7	83	9.1	63	US-10-060-425-10
8	81	8.9	1098	Sequence 41103, APP
9	80	8.8	162	Sequence 288, APP
10	80	8.8	177	Sequence 6604, APP
11	79	9	US-10-050-701-316	Sequence 316, APP
12	79	8.7	9	Sequence 48222, APP
13	79	8.7	US-10-194-975-101	Sequence 101, APP
14	77.5	8.7	9	Sequence 35, APP
15	77	8.4	119	Sequence 121, APP
16	77	8.4	150	Sequence 1674, APP
17	76.5	8.4	638	Sequence 4, APP
18	76.5	8.4	9	Sequence 38, APP
19	76.5	8.4	US-10-194-975-43	Sequence 43, APP
20	76.5	8.4	9	Sequence 3, APP

Post-processing: Minimum Match 0%, Maximum Match 100%

Listing first 45 summaries

ALIGNMENTS

RESULT 1

US-09-984-271-240

; Sequence 240, Application US/0984271

; Publication No. US20030040088A1

GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 71 Human Secreted Proteins

; FILE REFERENCE: PZ010P1

; CURRENT APPLICATION NUMBER: US/09/984,271

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 09/482,273

; PRIOR FILING DATE: 2000-01-13

; PRIOR APPLICATION NUMBER: PCT/US99/15849

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: 60/092,922

; PRIOR FILING DATE: 1998-07-15

; PRIOR APPLICATION NUMBER: 60/092,956

; PRIOR FILING DATE: 1998-07-15

; NUMBER OF SEQ ID NOS: 267

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 240

; LENGTH: 145

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-984-271-240

; Query Match Best Local Similarity 26.5%; Pred. No. 0.18; Score 93; DB 9; Length 145; Matches 36; Conservative 11; Mismatches 61; Indels 28; Gaps 5;

; Sequence 41103, A

; Sequence 288, APP

; Sequence 6604, APP

; Sequence 316, APP

; Sequence 48222, APP

; Sequence 101, APP

; Sequence 35, APP

; Sequence 121, APP

; Sequence 1674, APP

; Sequence 4, APP

; Sequence 38, APP

; Sequence 43, APP

; Sequence 3, APP

; Sequence 2, APP1

; Sequence 34, APP1

; Sequence 8, APP1

; Sequence 9, APP1

; Sequence 11, APP1

; Sequence 12, APP1

; Sequence 13, APP1

; Sequence 14, APP1

; Sequence 15, APP1

; Sequence 17, APP1

; Sequence 18, APP1

; Sequence 19, APP1

; Sequence 20, APP1

; Sequence 251, APP

; Sequence 1, APP1

; Sequence 3856, APP

; Sequence 1422, APP

; Sequence 5899, APP

; Sequence 5, APP1

; Sequence 42, APP1

; Sequence 52, APP1

; Sequence 53, APP1

Db 114 RGWAHCILCILVQL 129

; ORGANISM: homo sapiens
; US-10-217-774-4

RESULT 2
US-09-903-456-77
; Sequence 77, Application US/09903456
; Patient No. US2002013874A1
GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukeril, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
FILE REFERENCE: 6407.US.P3
CURRENT APPLICATION NUMBER: US/09/903,456
PRIORITY FILING DATE: 2001-07-11
PRIORITY APPLICATION NUMBER: US 09/624,670
PRIORITY FILING DATE: 2000-07-24
PRIORITY APPLICATION NUMBER: US 09/379,095
PRIORITY FILING DATE: 1999-08-23
SEQUENCE ID NOS: 116, 828
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 77
LENGTH: 265
TYPE: PRT
ORGANISM: Thraustochytrium aureum
US-09-903-456-77

Query Match 10.0%; Score 91; DB 10; Length 265;
Best Local Similarity 24.4%; Pred. No. 0.52; Mismatches 53; Conservative 16; Indels 60; Gaps 88; Gaps 10; SEQ ID NO 77

QY 16 PCW-LPWRATWWSS---STAWSS-----WASSA-----40
Db 40 PCWHSRWDTPWCCSSASRSSLRSRPLSSRPLSSCCTGCFSSDPCTCAWRPSARLSSEATK 99
QY 41 -LET-----STOPATGATWKWLHYAGSSRSRISPTLEATLTVSPFLA 80
Db 100 CLETWWRATSLMLRACLASCTCSCPCTSPRHTSSWIPPSRASRSPRSPCKT-TNPPFLP 153
QY 81 SLRVVARVCLR-----LICPPY-----PKDSSTEPSWRVAWSCPASLPAQLM 122
Db 159 SGGLSPSTLQEVMRTEFQSSTLSCPSCTHITSSPKGSQQSRTSPPRPS-----212
QY 123 SSPPRMPPTCLPVTKLTLRPWMAACGARVKERFLQTS 159
Db 213 ---WQCLCSPTTSSH---ATTTHRLLCSFLECTS 241

RESULT 3
US-10-217-774-4
; Sequence 4, Application US/10217774
; Publication No. US20020193583A1
GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: Same
; FILE REFERENCE: DEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774
; PRIORITY FILING DATE: 2002-08-12
; PRIORITY FILING DATE: 2001-08-14
; PRIORITY APPLICATION NUMBER: US 60/225,852
; PRIORITY FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT

RESULT 4
US-10-060-425-8

Query Match 9.6%; Score 88; DB 9; Length 1224;
Best Local Similarity 25.5%; Pred. No. 5.2; Mismatches 37; Conservative 18; Indels 56; Gaps 34; Gaps 7; Matches 37; APPLICANT: Hiebsch, Ronald
GENERAL INFORMATION:
; APPLICANT: Hiebsch, Ronald
TITLE OF INVENTION: Methods of Assessing Wolframin Protein Activity
FILE REFERENCE: 00450.US1
CURRENT APPLICATION NUMBER: US/10/060,425
PRIORITY FILING DATE: 2002-01-30
PRIORITY APPLICATION NUMBER: 60/266,385
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 8
LENGTH: 890
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-060-425-8

Query Match 9.6%; Score 87.5; DB 9; Length 890;
Best Local Similarity 21.2%; Pred. No. 4; Mismatches 28; Conservative 20; Indels 39; Gaps 5; Matches 28; APPLICANT: TWRLELAWEPCMMPWC-----LWRTWWSSSSSTAW-VWSASSALET 43
GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATTAS
; APPLICANT: GEORGE J. DAVISON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LERAY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUTIK
; APPLICANT: ISA K. MUSHAWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS

RESULT 5
US-08-424-550B-81

Query Match 9.6%; Score 88; DB 9; Length 1224;
Best Local Similarity 25.5%; Pred. No. 5.2; Mismatches 37; Conservative 18; Indels 56; Gaps 34; Gaps 7; Matches 37; APPLICANT: JOHN N. SIMONS
GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATTAS
; APPLICANT: GEORGE J. DAVISON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LERAY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUTIK
; APPLICANT: ISA K. MUSHAWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS

FEATURE: PRT
 ORGANISM: Homo sapiens
 OTHER INFORMATION: MAP TO AC022045.2
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8-15
 OTHER INFORMATION: EST_HUMAN HIT: BE008547.1, VALUE 6.00e-15

US-09-864-761-41103

Query Match 9.1%; Score 83; DB 10; Length 63;
 Best Local Similarity 38.7%; Pred. No. 0.57; Mismatches 25; Indels 10; Gaps 3;
 Matches 24; Conservative 3; Mismatches 25; Indels 10; Gaps 3;

Qy 93 CPPYKPSSTEPS-----WRYVAV-PSCPASLPAQIMSSRWWPRCPVLTLPWW 143
 Db 3 CRYPATPSSCWWSCPRYACHPWLRAWPPSCPSWSASFPSPSPP-TPCSSAASPSSLRSSW 61

Qy 144 AA 145
 Db 62 AS 63

RESULT 8
 US-09-712-363-289
 Sequence 288, Application US/09712363
 Patent No. US20020164588A1

GENERAL INFORMATION:
 APPLICANT: Eisenberg, David
 APPLICANT: Rotstein, Sergio H.
 APPLICANT: Malette, Edward M.

TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
 TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
 FILE REFERENCE: 07419-032001
 CURRENT APPLICATION NUMBER: US/09/712, 363
 CURRENT FILING DATE: 2000-11-13
 PRIOR APPLICATION NUMBER: PCT/US00/02246
 PRIOR FILING DATE: 2000-01-28
 PRIOR APPLICATION NUMBER: 60/179, 531
 PRIOR FILING DATE: 2000-02-01
 PRIOR APPLICATION NUMBER: 60/117, 844
 PRIOR FILING DATE: 1999-01-29
 PRIOR APPLICATION NUMBER: 60/118, 206,
 PRIOR FILING DATE: 1999-02-01
 PRIOR APPLICATION NUMBER: 60/126, 593
 PRIOR FILING DATE: 1999-03-26
 PRIOR APPLICATION NUMBER: 60/134, 093
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 60/134, 092
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 60/165, 124
 PRIOR FILING DATE: 1999-11-12
 PRIOR APPLICATION NUMBER: 60/165, 086
 NUMBER OF SEQ ID NOS: 292
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO: 288
 LENGTH: 1098
 TYPE: PRT
 ORGANISM: Mycobacterium tuberculosis
 US-09-712-363-288

Query Match 8.8%; Score 80; DB 9; Length 162;
 Best Local Similarity 28.9%; Pred. No. 3; Mismatches 44; Indels 17; Gaps 55;

Qy 18 WLPWNKWWNNSSSTAVNSASSALESTSTOPATGA--TWTKWLHYAGSSRISPTLRLTV 75
 Db 6 WLQRQSVPSAVSSPSSWSSSHOPWRNCPPHSAPSPSTHW---HSSRSP----- 54

Qy 76 SPFLASLRVARYCVLQLCPYPKPSSTESPRWVAMPSCPAS-----LPAOLMSSP 125
 Db 55 PPWCPT-----PCLDRPSF-SASALRPSASIRSRSRSPASSARSGHRSFSTEPSSLPLPSR 108

Qy 126 RW WFTCFLVTKLTLRPWWAACGARVKRRLQ 156
 Db 109 FWHRWEKC-----FTLQPAHAVT-KOIRRALQ 134

RESULT 10
 US-10-050-704-316
 Sequence 316, Application US/10050704
 Publication No. US20030050442A1

GENERAL INFORMATION:
 APPLICANT: Ruben et al.
 TITLE OF INVENTION: 62 Human Secreted Proteins
 FILE REFERENCE: P2039P1
 CURRENT APPLICATION NUMBER: US/10/050, 704
 CURRENT FILING DATE: 2003-01-18
 PRIOR APPLICATION NUMBER: 09/634, 524
 PRIOR FILING DATE: 2000-10-10
 PRIOR APPLICATION NUMBER: PCT/US00/089979
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/128, 693
 PRIOR FILING DATE: 1999-04-09
 PRIOR APPLICATION NUMBER: 60/130, 991

Qy 6 FLAWIPLCMMPCWLWPTTWMSSSSTAVNSASSALESTSTOP-----ANGATWTKWLH 58
 Db 628 FLAALFLFLALCWATNGWIVVSSGVPFNSAMPKIDGITSTIFFALPAHANGYAAWHL 687

RESULT 9
 US-09-738-626-6604
 Sequence 6604, Application US/09738626
 Publication No. US20020197605A1

GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738, 626
 CURRENT FILING DATE: 2000-12-19
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO: 6604
 LENGTH: 162
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-6604

Query Match 8.8%; Score 80; DB 9; Length 162;
 Best Local Similarity 28.9%; Pred. No. 3; Mismatches 44; Indels 17; Gaps 55;

Qy 18 WLPWNKWWNNSSSTAVNSASSALESTSTOPATGA--TWTKWLHYAGSSRISPTLRLTV 75
 Db 6 WLQRQSVPSAVSSPSSWSSSHOPWRNCPPHSAPSPSTHW---HSSRSP----- 54

Qy 76 SPFLASLRVARYCVLQLCPYPKPSSTESPRWVAMPSCPAS-----LPAOLMSSP 125
 Db 55 PPWCPT-----PCLDRPSF-SASALRPSASIRSRSRSPASSARSGHRSFSTEPSSLPLPSR 108

Qy 126 RW WFTCFLVTKLTLRPWWAACGARVKRRLQ 156
 Db 109 FWHRWEKC-----FTLQPAHAVT-KOIRRALQ 134

RESULT 10
 US-10-050-704-316
 Sequence 316, Application US/10050704
 Publication No. US20030050442A1

GENERAL INFORMATION:
 APPLICANT: Ruben et al.
 TITLE OF INVENTION: 62 Human Secreted Proteins
 FILE REFERENCE: P2039P1
 CURRENT APPLICATION NUMBER: US/10/050, 704
 CURRENT FILING DATE: 2003-01-18
 PRIOR APPLICATION NUMBER: 09/634, 524
 PRIOR FILING DATE: 2000-10-10
 PRIOR APPLICATION NUMBER: PCT/US00/089979
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/128, 693
 PRIOR FILING DATE: 1999-04-09
 PRIOR APPLICATION NUMBER: 60/130, 991

Qy 6 FLAWIPLCMMPCWLWPTTWMSSSSTAVNSASSALESTSTOP-----ANGATWTKWLH 58
 Db 628 FLAALFLFLALCWATNGWIVVSSGVPFNSAMPKIDGITSTIFFALPAHANGYAAWHL 687

RESULT 9
 US-09-738-626-6604
 Sequence 6604, Application US/09738626
 Publication No. US20020197605A1

GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738, 626
 CURRENT FILING DATE: 2000-12-19
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO: 6604
 LENGTH: 162
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-6604

Query Match 8.8%; Score 80; DB 9; Length 162;
 Best Local Similarity 28.9%; Pred. No. 3; Mismatches 44; Indels 17; Gaps 55;

Qy 18 WLPWNKWWNNSSSTAVNSASSALESTSTOPATGA--TWTKWLHYAGSSRISPTLRLTV 75
 Db 6 WLQRQSVPSAVSSPSSWSSSHOPWRNCPPHSAPSPSTHW---HSSRSP----- 54

Qy 76 SPFLASLRVARYCVLQLCPYPKPSSTESPRWVAMPSCPAS-----LPAOLMSSP 125
 Db 55 PPWCPT-----PCLDRPSF-SASALRPSASIRSRSRSPASSARSGHRSFSTEPSSLPLPSR 108

Qy 126 RW WFTCFLVTKLTLRPWWAACGARVKRRLQ 156
 Db 109 FWHRWEKC-----FTLQPAHAVT-KOIRRALQ 134

RESULT 10
 US-10-050-704-316
 Sequence 316, Application US/10050704
 Publication No. US20030050442A1

GENERAL INFORMATION:
 APPLICANT: Ruben et al.
 TITLE OF INVENTION: 62 Human Secreted Proteins
 FILE REFERENCE: P2039P1
 CURRENT APPLICATION NUMBER: US/10/050, 704
 CURRENT FILING DATE: 2003-01-18
 PRIOR APPLICATION NUMBER: 09/634, 524
 PRIOR FILING DATE: 2000-10-10
 PRIOR APPLICATION NUMBER: PCT/US00/089979
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/128, 693
 PRIOR FILING DATE: 1999-04-09
 PRIOR APPLICATION NUMBER: 60/130, 991

Qy 6 FLAWIPLCMMPCWLWPTTWMSSSSTAVNSASSALESTSTOP-----ANGATWTKWLH 58
 Db 628 FLAALFLFLALCWATNGWIVVSSGVPFNSAMPKIDGITSTIFFALPAHANGYAAWHL 687

PRIOR FILING DATE: 1999-04-26 ; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 344 ; PRIOR APPLICATION NUMBER: US 09/774, 203
; SOFTWARE: PatentIn Ver. 2.0 ; PRIOR FILING DATE: 2001-01-29
; SEQ ID NO 316 ; NUMBER OF SEQ ID NOS: 49117
; LENGTH: 177 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; TYPE: PRT ; SEQ ID NO 48222
; ORGANISM: Homo sapiens ; LENGTH: 90
; US-050-704-316 ; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; Query Match Best Local Similarity 8.8%; Score 80; DB 9; Length 177;
; Matches 33; Conservative 18; Mismatches 52; Indels 46; Gaps 7;
; QY 44 SPPGATGATW---TKWLYAGSSRISPTLETLTYSFP-----LASLRYA---RVC 88
; Db 1 SASCATGGSSWSRWTGTLGLTPRHRSQAATPLRASWWLATERWIMPGLCTTLELAWPSLRGC 60
; QY 89 LRLCPYPKDSSTEPSWRVAVPSC-----PASLPALMSSPRRWPF-- 130
; Db 61 SLPFACTVLSPTRKGPPRWTWLPCVWLSSPLSPWSSVSSLSMRVLSCHMGQPCVG 120
; QY 131 CL-PVTKILRDP-----WWIAC 146
; Db 121 CYSISSESMAPSAATSGQSPOTHWHLIC 149

RESULT 11
US-09-864-761-48222
; Sequence 48222, Application US/09864761
; Patent No. US2002008763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aegomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180, 312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207, 456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632, 366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236, 359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234, 687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608, 408

RESULT 12
US-10-194-975-101
; Sequence 101, Application US/10194975
; Publication No. US2003039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231_01
; CURRENT APPLICATION NUMBER: US/10/194, 975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305, 111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122 ; SEQ ID NO 101
; LENGTH: 96
; SOFTWARE: PatentIn version 3.1
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-194-975-101
; Query Match Best Local Similarity 8.7%; Score 79; DB 9; Length 96;
; Matches 17; Conservative 11; Mismatches 14; Indels 8; Gaps 2;
; QY 27 SSSTAAWWSASSALETSTQPATGATWPKWLYAGSSISPTLETLTYS 76
; Db 28 SISSYYW-SWI-----RQPPKGLEWIGIYIYSGSTNNPSLKSRYVIS 69

RESULT 13
US-10-194-975-35
; Sequence 35, Application US/10194975
; Publication No. US2003039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 50131_01
; CURRENT APPLICATION NUMBER: US/10/194, 975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305, 111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Version 3.1

```

; SEQ ID NO 35 ; SEQ ID NO 1674
; LENGTH: 98 ; LENGTH: 150
; TYPE: PRT ; TYPE: PRT
; ORGANISM: Homo sapiens ; ORGANISM: Homo sapiens
; US-10-194-975-35 ; FEATURE:
;                                     NAME/KEY: VARIANT
;                                     LOCATION: (55)
;                                     OTHER INFORMATION: Wherein Xaa may be any one of Ala or Arg or Gln or Glu or Gly
;                                     OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
;                                     OTHER INFORMATION: US-09-867-550-1674
;                                     OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
;                                     OTHER INFORMATION: US-09-867-550-1674

Query Match 8.7%; Score 79; DB 9; Length 98;
Best Local Similarity 29.4%; Pred. No. 2.1; 15; Mismatches 13; Indels 17; Gaps 6; Gaps 1;
Matches 15; Conservative 13; Mismatches 17; Indels 6; Gaps 1;

OY 26 WSSSTAWSWASAALETSTOPATGATTWKWHYAGSRISPTLATVTS 76
Db 27 YSSSSNWWGWIRQ----PPGGLEWIGYSGSTVYNPSLKSRVTMS 71

RESULT 14
US-09-860-670-121
; Sequence 121, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION DATA REMOVED - CONSULT PALM OR FILE WRAPPER
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (102)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (119)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-860-670-121

Query Match 8.4%; Score 77; DB 10; Length 150;
Best Local Similarity 29.4%; Pred. No. 5.1; 13; Mismatches 38; Indels 50; Gaps 10;
Matches 43; Conservative 13; Mismatches 38; Indels 50; Gaps 10;

OY 26 WSSSTAWSWASAALETSTOPATGATTWKWHYAGSRISPTLATVTS 82
Db 15 WSSEATP---APSAYPSSCRARAAFFTSKARFA---SP-----VSGLXAAR 58
;                                     NAME/KEY: VARIANT
;                                     LOCATION: (55)
;                                     OTHER INFORMATION: Wherein Xaa may be any one of Ala or Arg or Gln or Glu or Gly
;                                     OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
;                                     OTHER INFORMATION: US-09-867-550-1674
;                                     OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
;                                     OTHER INFORMATION: US-09-867-550-1674

Query Match 8.4%; Score 77; DB 10; Length 150;
Best Local Similarity 29.4%; Pred. No. 5.1; 13; Mismatches 38; Indels 50; Gaps 10;
Matches 43; Conservative 13; Mismatches 38; Indels 50; Gaps 10;

OY 83 RVAR-----VCLR-----LCPYPKDSSTEPSWR-----VAMPSCASLPAQL----- 121
Db 59 RGGRSLPGRSSCPRLGTAVCTP-----DPWRGGALRTGGPLDORHRRPPPP 111
;                                     NAME/KEY: VARIANT
;                                     LOCATION: (55)
;                                     OTHER INFORMATION: Wherein Xaa may be any one of Ala or Arg or Gln or Glu or Gly
;                                     OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
;                                     OTHER INFORMATION: US-09-860-670-121
;                                     OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
;                                     OTHER INFORMATION: US-09-860-670-121

OY 122 -MSSRWWPTCLPVKLRPWW 143
Db 112 PGISSRLPPVAMPSGLSSSWW 134

Search completed: March 27, 2003, 16:20:25
Job time: 39.8421 secs

```

RESULT 15

US-09-867-550-1674

; Sequence 1674, Application US/09867550

; Patent No. US20020165137A1

; GENERAL INFORMATION:

; APPLICANT: Leach, Martin D.

; APPLICANT: Mehraban, Fuad

; APPLICANT: Conley, Pamela

; APPLICANT: Law, Debbie

; APPLICANT: Topper, James

TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and

FILE REFERENCE: 21A03-013 (Cur-313)

CURRENT APPLICATION NUMBER: US/09/867,550

PRIOR APPLICATION NUMBER: USN 60/208,427

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FASTSEQ FOR WINDOWS Version 4.0

GenCore version 5.1.4_p5_4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:01:27 ; Search time 27.4737 seconds

(without alignments)
 566.862 Million cell updates/sec

Title: US-10-019-219-1
 Perfect score: 912
 Sequence: TIVRFLAWLPCMMTPCWLP.....WAACGARVKRFLQLTLSR 162

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing first 45 summaries
 Database : PIR_7;*: 1: pir1;* 2: pir2;* 3: pir3;* 4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1

S73704 hypothetical protein H08_orf157a - Mycoplasma pneumoniae (strain ATCC 29342)

C;Species: Mycoplasma pneumoniae
 A;Variety: ATCC 29342
 C;Date: 24-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000

C;Accession: S73704 R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996

A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
 A;Reference number: S73327; MNUID:97105685; PMID:6948633

A;Accession: S73704 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-157 <HIM>

A;Cross-references: EMBL:AE000336; GB:U00059; NIRD:91674053; PRON:AB96025.1; PRD:167

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C;Genetics:
 A;Genetic code: SGC3

C;Superfamily: Mycoplasma pneumoniae hypothetical protein H08_orf157a

Query	Match	Length	DB ID	Description
Qy	9 WLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	7 332 2 C72310	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	8 9.1 F95948	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	9 8.5 B83286	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	10 8.9 T03265	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	11 8.1 G70697	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	12 8.9 1522 2 S35049	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	13 8.0 610 2 JMW067	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	14 7.5 660 2 S26903	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	15 7.9 98 2 S12414	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	16 7.9 98 2 S12414	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	17 7.9 1007 2 T01437	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	18 7.6 784 2 JQ0317	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	19 7.6 3570 2 T45025	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	20 7.5 253 2 AC0810	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	21 7.5 263 1 AC43199	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	22 7.7 118 2 A72474	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	23 7.7 145 2 S78055	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	24 7.7 146 2 S09711	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	25 7.7 211 2 S73791	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	26 7.7 1513 2 A54895	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	27 7.7 1859 1 S74293	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	28 7.7 1862 2 T29599	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----
Qy	29 7.7 512412	496	Db	WMLPCMMTPCWLP-----W-KTWWSSSSTAVWS-----ASSALETSTQ-----

Ig heavy chain pre Ig heavy chain Vr
 Ig heavy chain - Sal cysteine - Sal
 rnf protein - Rho malate dehydrogena N-methyl-D-asparta probable mucin DKF hypothetical prote valyl-tRNA synthet valyl-tRNA synthet protein-tirosine-p Ig heavy chain Vr hypothetical prote hypothetical prote flt3 ligand - huma paired-type homeod

A;Cross-references: EMBL:53341; PIDN:AAE9106_1; GSPDB:GN00028; CESP:F49E10.2
A;Experimental source: strain Bristol N2; clone F49E10
C;Genetics:
A;Gene: CES:F49E10.2
A;Map position: X
A;Introns: 51/1; 92/3; 176/3; 235/3; 332/2; 514/1; 543/2; 569/3; 677/1; 732/3
Query Match 9.6%; Score 87.5; DB 2; Length 790;
Best Local Similarity 24.8%; Pred. No. 2, 9; Mismatches 39; Indels 39; Gaps 4;
Matches 29; Conservative 10; Mismatches 39; Indels 39; Gaps 4;

Oy 27 SSSSTAWWSWASSALETSQPATGATWKWLHAGGSRSRISPTLEATIVSPLASRVAR 86
Db 627 TSTSTT^{TTT}TATTTT^{TTT}TSEK-----TPPTTIPKIPRWPW--LAGSSSTEQPHW 704

Oy 87 VCLRLLCPPYKDSSEPSWRVAVPSCASPAPQLMSSPRWPCILPVTKLTLRPW 143
Db 667 -----PPTTIPKIPRWPW-----TPPTTIPKIPRWPW--LAGSSSTEQPHW 704

RESULT 3
JQ0137 hypothetical 30.1K protein - *Pseudomonas aeruginosa*
C;Species: *Pseudomonas aeruginosa*
C;Accession: JQ0137
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 07-Jun-1996
R;Kato, J.; Chu, L.; Kitano, K.; Default, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra, T.
Gen. 84, 31-38, 1989
A;Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in *Ps*
A;Reference number: JQ0132; MID:90108714; PMID:2514124
A;Accession: JQ0137
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-261 <RAW>
A;Note: 3' Met could also be the initiator
C;Genetics:
A;Start codon: GTG

Query Match 9.4%; Score 85.5; DB 2; Length 261;
Best Local Similarity 30.0%; Pred. No. 1, 4; Mismatches 8; Indels 30; Gaps 7;
Matches 33; Conservative 8; Mismatches 30; Indels 39; Gaps 7;

Oy 44 STQP^TGATWKWLHAGGSRSRISPTLEATIVSPLASRVARVCLRLCPPYKDSSE 103
Db 64 SPQR^SQERWRRAWLVRQSRVSP-----QS 99

Oy 104 PSW-----RV---AVPSCPASLPA-OLMSSPRWWPTCLPVTKLTLRPW 143
Db 100 PAW^SQASRPPVSPHAWP---PAWLRSRLRSRSPRAWP---IVSPDASPPW 144

RESULT 4
H72460 hypothetical protein APE2332 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Accession: H72460
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
R;Kavarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaya, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*
A;Reference number: A72450; MID:99310339; PMID:10382966
A;Accession: H72460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-210 <RAW>
A;Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BA81344_1; PID:d1045130; RID:95105945; AP000064
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2332

Query Match 9.2%; Score 83.5; DB 2; Length 384;
Best Local Similarity 25.4%; Pred. No. 3, 2; Mismatches 36; Conservative 8; Mismatches 51; Indels 47; Gaps 8; Matches 36; Conservative 8; Mismatches 51; Indels 47; Gaps 8;

Oy 25 WWSSTAWWSWASSALETSQPATGATWKWLH-----YAG--SSRISPTLE 70
Db 12 WWS^S-----LRW----FGVSIAPGVGGFWHWWQSLAWLPMRCRQVGMGLSERILLSPQ 62

Oy 71 ATLTVPFLSRLVARVCLRLCPPYKDSSEPSWRVAVPSCASPAPQLMSS---PR 126
Db 63 AD-----GLHLVYRQCGVLU-----EPVQVPRPITQOELSGNLPIQIPLR 104

Oy 127 WWP^TCLPVTKLTLRPWWAACGA 148
Db 105 HW-LLPAT^SALCRPLRLPAGA 124

RESULT 5
G82670 general secretory pathway protein L XFL1524 [imported] - *Xylella fastidiosa* (strain 9a)
C;Species: *Xylella fastidiosa*
C;Accession: G82670
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
R;Anonymous: The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Seq
A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A;Reference number: A82515; MID:20565717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82670
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <SIM>
A;Cross-references: GB:AE003982; GB:AE003849; NID:99106554; PIDN:AAF84333_1; GSPDB:GN
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Aceencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carrasco, D.M.; Carrer
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Farro, J.A.; Froga, J.S.; Franca, S.C.; Franco, M.C.; Fr
chado, M.R.; Jungueira, M.L.; Kempfer, E.L.; Krajcmar, J.P.; Krieger, J.E.; Kurama, E.E.; La
do, M.R.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.J.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsuura, A.Y.; Menck, C.F.M.; Mircca, E.C.; Miyake, C.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva Jr., W.A.; da Silva
M.; Tsuhako, M.H.; Vallada, H.; van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XFL1524

Query Match 9.2%; Score 83.5; DB 2; Length 384;
Best Local Similarity 25.4%; Pred. No. 3, 2; Mismatches 36; Conservative 8; Mismatches 51; Indels 47; Gaps 8; Matches 36; Conservative 8; Mismatches 51; Indels 47; Gaps 8;

Oy 25 WWSSTAWWSWASSALETSQPATGATWKWLH-----YAG--SSRISPTLE 70
Db 12 WWS^S-----LRW----FGVSIAPGVGGFWHWWQSLAWLPMRCRQVGMGLSERILLSPQ 62

Oy 71 ATLTVPFLSRLVARVCLRLCPPYKDSSEPSWRVAVPSCASPAPQLMSS---PR 126
Db 63 AD-----GLHLVYRQCGVLU-----EPVQVPRPITQOELSGNLPIQIPLR 104

Oy 127 WWP^TCLPVTKLTLRPWWAACGA 148
Db 105 HW-LLPAT^SALCRPLRLPAGA 124

RESULT 6
A54770 N-acetylglucosamine-specific receptor 1 precursor - human
C;Species: Homo sapiens (man)
C;Accession: A54770; S37024
R;Blanc, O.; Perrin, C.; Mziaut, H.; Darbon, H.; Mattei, M.G.; Miquelis, R.

Query Match Similarity 9.3%; Score 84.5; DB 2; Length 210;
Best Local Similarity 25.0%; Pred. No. 1, 4; Length 210;

Genomics 21, 18-26, 1994
A;Title: Molecular cloning, cDNA analysis, and localization of a monomer of the N_{acetyl} Glucosaminidase from *Neurospora crassa*
A;Reference number: A54770; MUID:94375011; PMID:8088785
A;Accession: A54770
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-496 <BLA>
A;Cross-references: EMBL:X72018
C;Genetics:
A;Gene: GDB:NAGR1
A;Cross-references: GDB:250465; OMIM:160994
A;Sap position: 19p13.2-19p13.2
C;Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C;Keywords: calcium binding; lectin; thyroid gland; transmembrane protein
F;1-187/domain: signal sequence #status predicted <SIQ>
F;19-496/product: N-acetylglucosamine receptor 1 #status Predicted <MAT>
F;497-511/domain: ribonucleoprotein repeat homology <RRM2>
F;530-536/Region: glycine-rich
F;441-460/Domain: transmembrane #status Predicted <TM>
F;488-495/Region: coated-pit mediated internalization signal

Query Match 9.2%; Score 83.5; DB 2; Length 496;
Best Local Similarity 23.7%; Pred. No. 4.2; Mismatches 13; Indels 39; Gaps 6;
Matches 27; Conservative 13; Mismatches 35; Indels 39; Gaps 6;

Qy 14 MVCWLPLRTWMMWSSTAWNSWAALSTIQTPTGATWTKLHYRGSSRSITPLETL 73
Db 391 MPPAWSAWAPTIW-SGNAWSAWAPTA-----SSAWNSAWAPTAASSAWAP----- 435

Qy 74 TVSPFLASLVARVAVCLRLJCQCPYKPKDSSTEPSRNSRVAVWSPCPASPAQLMSSPRW 127
Db 436 WARFWAL-----SAW--AWPWWAVAVPALVPS-RW 465

RESULT 7
C72310
conserved hypothetical protein - *Thermotoga maritima* (strain MSB8)
C;Species: *Thermotoga maritima*
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: C72310
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, M.M.; Stewart, A.M.; Cottrell, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; Garrett, C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between archaea and bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: C72310
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <ARN>
A;Cross-references: GB:AE001760; GB:AE000512; NID:94981510; PIDN:AAD36061.1; PID:9498152
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0982

Query Match 9.1%; Score 83; DB 2; Length 332;
Best Local Similarity 26.8%; Pred. No. 3.1; Mismatches 30; Conservative 20; Indels 42; Gaps 6;

Qy 30 STRAWSWASSALER-STPATGATWTKWLH-----YAGSSRSRPLTLETLIVS 76
Db 104 TAMEAAALAYALTAAHTAKKGFLSEWIKWMSNTYTVTSNTNPVY--PKGKPTIAVLG 161

Qy 77 PFLASLRVARVCLRLJCQCPYKPKDSSTEPSRNSRVAVWSPCPASPAQLMSSPRW 128
Db 162 PWIAS-TIFLAMMIWYAGVKNKSQRPS-KLNW--IAASVILIAPIWW 208

RESULT 8
F95948
probable acyltransferase, possibly surface-saccharide specific acetyltransferase protein
C;Species: *Sinorhizobium meliloti*
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C;Accession: F95948
R;Final, T.M.; Weindner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb PSymB megaplasmid from the N2-fixing e
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: F95948
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-377 <KUR>
A;Cross-references: GB:AL591985; PIDN:GAC49254.1; PID:915140740; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid PSymB
R;Galibert, F.; Flian, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lealau
heault, P.; Vandenberg, M.; Vorholter, F.J.; Weindner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Smb2118
A;Genome: plasmid
Query Match 9.1%; Score 83; DB 2; Length 377;
Best Local Similarity 22.8%; Pred. No. 3.5; Mismatches 36; Conservative 11; Indels 54; Gaps 7;
Matches 36; Conservative 11; Mismatches 57; Indels 54; Gaps 7;

Qy 18 WLPW--RWMMWSSTAWNSWAALSTIQTPTGATWTKLHYAGSSRSITPLETL 171
Db 120 WTPWVLTCHTLSLSDEQFYLVWPLIVLVPRSSVAGVCVGIVCSLAYRHYW-----WTKLHYAG 61

Qy 62 SSRISPTLETLIVSPLFPLSLRVARVAVCLRLJCQCPYKPKDS-----STPSWRVAPWSCP 114
Db 172 -----PLTGCPSLA-----RDLLPPASMDALAVGALLAARPSWRSGWPAWA 212

Qy 115 --ASIPQAQMSSPRQPTCLPVTKLTLPRLPWAAGARY 150
Db 213 KLSWMPLSLASLCLVWSK--PVAMPVVAFWAGILEV 248

RESULT 9
B83286
hypothetical protein PA2884 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C;Species: *Pseudomonas aeruginosa*
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: B83286
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
admin, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Harbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: B83286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-254 <STO>
A;Cross-references: GB:AE004714; GB:AE004091; NID:9948965; PIDN:AAG06272.1; GSPDB:GN
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2884
Query Match 8.9%; Score 81.5; DB 2; Length 254;
Best Local Similarity 23.1%; Pred. No. 3.2; Mismatches 37; Conservative 14; Indels 62; Gaps 47; Gaps 7;

Qy 2 VRLFLAWLPCMMVCPWLPDPMWNSSTAWNSWAAL-----ET 43
Db 30 IVLYVIW--TYAERWLPLW-----PVSLAASIAIAHLPGFPLRLFPWDLA 80

Qy 44 STOPANGATWTKWLHYAGSSRSRPLTLETLIVSPP--LASLRARV--CLRLJCQCPYKPD 99
Db 81 SASPP1QMIWLPW---DKALVALTLLAWLRLRPQQLVSDLTIALAFCUFFFVPLLSI 136

RESULT 10

QY 100 SSTEPPWRRVAMPSCPASLPAQLMSSPPIKPLPVVKLT 139
 C;Species: Homo sapiens (man)
 C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
 C;Accession: T00326
 R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
 DNA Res. 5, 31-39, 1998
 A;Title: Prediction of the coding sequences of unidentified human genes. IX. The completon
 A;Reference number: 214086; MUID:98290545; PMID:9628581
 A;Accession: T00326
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: mRNA
 A;Residues: 1-984 <NAG>
 A;Cross-references: EMBL:AB011122; NID:93043623; PIDN:BAA25476.1; PID:93043624
 A;Experimental source: brain
 C;Genetics:
 C;Note: KIAA0550
 C;Superfamily: thrombospondin type 1 repeat homology <THR3>
 F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 8.9%; Score 81; DB 2; Length 984;
 Best Local Similarity 21.6%; Pred. No. 14; Mismatches 59; Indels 44; Gaps 6;
 Matches 32; Conservative 13; Mismatches 59; Indels 44; Gaps 6;

QY 11 PCMMWPC----WLPNPTWWNS-----SSPRAVWRASSALEST 45
 Db 391 PCNIALCPVDGQWOENSSWWSQCSCVICSNGTQORSROCTAAHGGSECGRPWAESRECYNP 450

QY 46 QPATGATWKWLHYAGSS-----RISPLEATLVSPFLASLRVARVCLRLCP-P 95
 Db 451 ECTANGQMNQWHGWSCKSCDGGWERRIRTCQGAVITGQQCEGTGBEVRRCESEQRCAP 510

QY 96 Y---PKDSSTEPSWR----VAWPSCP 114
 Db 511 YEICPEDYLMMSMWKRTPAGDLAFNQCP 538

RESULT 11

QY 59 YAGSSRISPLTEATIVSP----FLASLRVARVCLRULCPIPYPKDSS 101
 C;Species: Homo sapiens (man)
 C;Accession: T00028
 R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
 Cytoogenet. Cell Genet. 79, 103-108, 1997
 A;Title: Cloning and characterization of BA12 and BA13, novel genes homologous to brain-specific angiogenesis inhibitor 3 - human
 A;Reference number: 21066; MUID:9194217; PMID:9533023
 A;Accession: T00028
 A;Status: translated from GB/EMBL/DDJB
 A;Molecule type: mRNA
 A;Residues: 1-1522 <SH1>
 A;Cross-references: EMBL:AB005299; NID:93021700; PIDN:BAA25363.1; PID:93021701
 A;Experimental source: brain
 C;Genetics:
 A;Gene: GDB:BA13
 A;Cross-references: GDB:9838090; OMIM:602684
 A;Map position: 6q12-6q12
 C;Superfamily: thrombospondin type 1 repeat homology <THR3>

Query Match 8.9%; Score 81; DB 2; Length 1522;
 Best Local Similarity 21.6%; Pred. No. 22; Mismatches 59; Indels 44; Gaps 6;
 Matches 32; Conservative 13; Mismatches 59; Indels 44; Gaps 6;

QY 11 PCMMWPC----WLPNPTWWNS-----SSTANVWRASSALEST 45
 Db 391 PCNIALCPVDGQWOENSSWWSQCSCVICSNGTQORSROCTAAHGGSECGRPWAESRECYNP 450

QY 46 QPATGATWKWLHYAGSS-----RISPLEATLVSPFLASLRVARVCLRLCP-P 95
 Db 451 ECTANGQMNQWHGWSCKSCDGGWERRIRTCQGAVITGQQCEGTGBEVRRCESEQRCAP 510

QY 96 Y---PKDSSTEPSWR----VAWPSCP 114
 Db 511 YEICPEDYLMMSMWKRTPAGDLAFNQCP 538

RESULT 12

QY 59 YAGSSRISPLTEATIVSP----FLASLRVARVCLRULCPIPYPKDSS 101
 C;Species: Homo sapiens (man)
 C;Accession: T00028
 R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
 Cytoogenet. Cell Genet. 79, 103-108, 1997
 A;Title: Cloning and characterization of BA12 and BA13, novel genes homologous to brain-specific angiogenesis inhibitor 3 - human
 A;Reference number: 21066; MUID:9194217; PMID:9533023
 A;Accession: T00028
 A;Status: translated from GB/EMBL/DDJB
 A;Molecule type: mRNA
 A;Residues: 1-1522 <SH1>
 A;Cross-references: EMBL:AB005299; NID:93021700; PIDN:BAA25363.1; PID:93021701
 A;Experimental source: brain
 C;Genetics:
 A;Gene: GDB:BA13
 A;Cross-references: GDB:9838090; OMIM:602684
 A;Map position: 6q12-6q12
 C;Superfamily: thrombospondin type 1 repeat homology <THR3>

Query Match 8.9%; Score 81; DB 2; Length 1522;
 Best Local Similarity 21.6%; Pred. No. 22; Mismatches 59; Indels 44; Gaps 6;
 Matches 32; Conservative 13; Mismatches 59; Indels 44; Gaps 6;

QY 11 PCMMWPC----WLPNPTWWNS-----SSTANVWRASSALEST 45
 Db 391 PCNIALCPVDGQWOENSSWWSQCSCVICSNGTQORSROCTAAHGGSECGRPWAESRECYNP 450

QY 46 QPATGATWKWLHYAGSS-----RISPLEATLVSPFLASLRVARVCLRLCP-P 95
 Db 451 ECTANGQMNQWHGWSCKSCDGGWERRIRTCQGAVITGQQCEGTGBEVRRCESEQRCAP 510

QY 96 Y---PKDSSTEPSWR----VAWPSCP 114
 Db 511 YEICPEDYLMMSMWKRTPAGDLAFNQCP 538

RESULT 13

QY 59 YAGSSRISPLTEATIVSP----FLASLRVARVCLRULCPIPYPKDSS 101
 C;Species: Homo sapiens (man)
 C;Accession: T00028
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davis, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:9829587; PMID:9634230
 A;Accession: T00697
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1098 <COL>
 A;Cross-references: GB:260343; GB:AU123456; NID:93261648; PIDN:CAB02474.1; PID:91552878
 C;Genetics:
 A;Gene: embB

Query Match 8.9%; Score 81; DB 2; Length 1098;
 Best Local Similarity 26.8%; Pred. No. 16; Mismatches 50; Indels 14; Gaps 3;

QY 6 FLAWIIPCMMPWCWLWPTTWWSSSSTAWVSRAASSALETSTQP-----ATGATWKWLH 58
 Db 628 FLAALFLFLALCWAATGWWVWVSSIGVPFNSAMPKIDGITVSTIFALFAAGAYWHLH 687

RESULT 14

QY 59 YAGSSRISPLTEATIVSP----FLASLRVARVCLRULCPIPYPKDSS 101
 C;Species: Homo sapiens (man)
 C;Accession: T00028
 R;Aubert, J.
 A;Title: Submission to the EMBL Data Library, September 1993
 A;Reference number: S37593
 A;Accession: S37594
 A;Molecule type: mRNA
 A;Residues: 1-20, 'W', 22-610 <AUB>
 A;Cross-references: EMBL:X/4955

Query Match 8.8%; Score 80; DB 2; Length 610;
 Best Local Similarity 23.9%; Pred. No. 11; Mismatches 16; Conservative 16; Mismatches 19; Indels 16; Gaps 2;

THIS PAGE BLANK (USPTO)

DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Wolframin.
 GN WFS1.

OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutetrapoda;
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 OX NCBI_TAXID=10090;
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=99036670; PubMed=9817917;

RA Strom T.M., Hoertnagel K., Hofmann S., Gekeler F., Scharfe C.,
 Rabi W., Gerbitz K.-D., Meitinger T.;
 RT "Diabetes insipidus, diabetes mellitus, optic atrophy and deafness
 (DIDMOAD) caused by mutations in a novel gene (Wolframin) coding for
 a predicted transmembrane protein.";
 RT Hum. Mol. Genet. 7:2021-2028(1998).

[2]

RN SEQUENCE FROM N.A.

RX TISSUE=insulinoma;
 RX MEDLINE=98442649; PubMed=9771706;

RA Inoue H., Tanizawa Y., Wasson J., Behn P., Kalidas K.,
 RA Bernal-Mirachchi E., Mueckler M., Marshall H., Donis-Keller H.,
 RA Crook P., Rogers D., Mikuni M., Kumashiro H., Higashi K., Sobue G.,
 RA Oka Y., Permutt M.A.;

CC "A gene encoding a transmembrane protein is mutated in patients with
 diabetes mellitus and optic atrophy (Wolframin Syndrome).";

CC Nat. Genet. 20:143-148(1998);

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 reticulum.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AJ011971; CAA09892.1. :-

DR AF084482; AAC64944.1. :-

DR MGD; MGI:1328355; Wfs1.

KW transmembrane.

FT TRANSMEM 314 POTENTIAL.
 FT TRANSMEM 340 POTENTIAL.
 FT TRANSMEM 402 POTENTIAL.
 FT TRANSMEM 427 POTENTIAL.
 FT TRANSMEM 447 POTENTIAL.
 FT TRANSMEM 455 POTENTIAL.
 FT TRANSMEM 496 POTENTIAL.
 FT TRANSMEM 529 POTENTIAL.
 FT TRANSMEM 563 POTENTIAL.
 FT TRANSMEM 589 POTENTIAL.
 FT TRANSMEM 632 POTENTIAL.
 FT TRANSMEM 870 POTENTIAL.
 FT DOMAIN 13 POLY-LYS.
 FT DOMAIN 191 POLY-LYS.
 FT CONFLICT 215 A -> V (IN REF. 2).
 FT SEQUENCE 890 AA; 100578 MW; CB6GCA16171a942 CRC64;

Query Match 9.4%; Score 85.5; DB 1; Length 890;
 Best Local Similarity 21.2%; Pred. No. 2.5;
 Matches 28; Conservative 20; Mismatches 45; Indels 39; Gaps 5;

QY 1 TVVRFLFLAWLPCMMVPCW-----LPARTWWSSSPAW-VSWASSALET 43
 Db 633 SMVKLILVWETAILFCWFTYVYRSEGKVNSTLIWQQYGLCGRRAWKETNMARQILC 692

Qy 44 STQPATGATWPKWLVHAGSSRISPTLEATIVSPLAS--LRVARVCLRLCPPIPDKSS 101
 Db 693 SHLEGHRVWTGFRKFYVVRVTEIDNSAESAESAINMLPFLGDNMR---CL----- 736

RESULT 3

EMBL_MCTU : 1:1| ID EMBL_MCTU STANDARD; PRT; 1098 AA.
 AC P72030; P72061;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DR 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable arabinosyltransferase B (EC 2.4.2.-).
 GN EMBB OR RV3795 OR MTCV13D12.29.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinomycetales; Actinomycetaceae; Actinomycetales; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TAXID=1773;
 OX NCBI_TAXID=1773;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37Rv;

RX MEDLINE=97287037; PubMed=9142129;

RA Telenti A., Philipp W.J., Sreevatsan S., Bernasconi C.,
 RA Stockbauer K.E., Wieles B., Musser J.M., Jacobs W.R. Jr.;
 RT "The emb operon, a gene cluster of Mycobacterium tuberculosis involved
 in resistance to ethambutol.;"
 RT Nat. Med. 3:567-570(1997).

[2]

RN SEQUENCE FROM N.A.

RC STRAIN=H37Rv;

RX MEDLINE=98285987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Barry C.E. III, Tekla F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moul S., Murphy L.,
 RA Oliver S., Osborne J., Quill M.A., Rajandream M.A., Roberts J.,
 RA Rutter S., Seeger K., Skelton M., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.;"
 RT Nature 393:537-544(1998).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwynn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Emiliaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J.J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains;"
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RR VARIANTS EMB RESISTANT L306; I-306; V-306 AND V-330.
 RX MEDLINE=9740246; PubMed=9257740;
 RA Sreevatsan S., Stockbauer K.E., Pan X., Kreiswirth B.N.,
 RA Moghazeh S.I., Jacobs W.R. Jr., Telenti A., Musser J.M.,
 RT "Ethambutol resistance in Mycobacterium tuberculosis: critical role of
 RT embB mutations.;"
 RT Antimicrob. Agents Chemother. 41:1677-1681(1997).

RN [5]

RP VARIANTS EMB RESISTANT.

RX MEDLINE=20106977; PubMed=10639358;

RA Ramasamy S.V., Amin A.G., Goeksel S., Stager C.E., Dou S.-J.,
 RA El-Sawy H., Moghazeh S.I., Kreiswirth B.N., Musser J.M.;
 RT "Molecular genetic analysis of nucleotide polymorphisms associated
 RT with ethambutol resistance in human isolates of Mycobacterium
 RT tuberculosis.;"
 RT Antimicrob. Agents Chemother. 44:326-336(2000).

RN [6]

RP VARIANTS EMB RESISTANT LEU-306; ILE-306 AND VAL-306.
 RX MEDLINE=21106283; PubMed=11162078;

RA Rinder H., Mieskes K.T., Tortoli E., Richter E., Casal M., Vaquero M., RA Cambau E., Feldmann K., Loscher T.;
 RT "Detection of embB codon 306 mutations in ethambutol resistant
 RT Mycobacterium tuberculosis directly from sputum samples: a low-cost,
 RL rapid approach"; Mol. Cell. Probes 15:37-42(2001).

CC -1- FUNCTION: Arabinosyl transferase responsible for the
 CC polymerization of arabinos into the arabinan of arabinogalactan.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- MISCELLANEOUS: This is one of the target of the anti-tuberculosis
 CC drug ethambutol ((S,S')-2,2'-(ethylenedimino)dI-1-butanol; EMB).
 CC EMB is a first-line drug used to treat tuberculosis. EMB inhibits
 CC the transfer of arabinogalactan into the cell wall.
 CC -1- SIMILARITY: BELONGS TO THE EMB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: 280343; CAB02474.1; -
 DR EMBL: U68480; AAC5281.1; -
 DR EMBL: AB007183; AAC5268.1; -
 DR TIGR: MM3902; -
 DR DRUG: Glycosyltransferase; Transmembrane; Cell wall;
 KW Antibiotic resistance; Complete proteome.
 FT TRANSMEM 28 50 POTENTIAL.
 FT TRANSMEM 217 239 POTENTIAL.
 FT TRANSMEM 271 293 POTENTIAL.
 FT TRANSMEM 402 419 POTENTIAL.
 FT TRANSMEM 434 456 POTENTIAL.
 FT TRANSMEM 472 494 POTENTIAL.
 FT TRANSMEM 541 558 POTENTIAL.
 FT TRANSMEM 570 587 POTENTIAL.
 FT TRANSMEM 597 619 POTENTIAL.
 FT TRANSMEM 626 648 POTENTIAL.
 FT TRANSMEM 663 685 POTENTIAL.
 FT TRANSMEM 698 720 POTENTIAL.
 FT VARIANT 297 297 S -> A (RESISTANCE TO EMB).
 FT VARIANT 306 306 M -> I (RESISTANCE TO EMB).
 FT VARIANT 306 306 M -> L (RESISTANCE TO EMB).
 FT VARIANT 306 306 M -> V (RESISTANCE TO EMB).
 FT VARIANT 328 328 D -> G (RESISTANCE TO EMB).
 FT VARIANT 328 328 D -> Y (RESISTANCE TO EMB).
 FT VARIANT 330 330 F -> V (RESISTANCE TO EMB).
 FT VARIANT 334 334 Y -> H (RESISTANCE TO EMB).
 FT VARIANT 406 406 G -> A (RESISTANCE TO EMB).
 FT VARIANT 406 406 G -> C (RESISTANCE TO EMB).
 FT VARIANT 406 406 G -> D (RESISTANCE TO EMB).
 FT VARIANT 497 497 Q -> K (RESISTANCE TO EMB).
 FT VARIANT 497 497 Q -> R (RESISTANCE TO EMB).
 FT VARIANT 745 745 G -> D (RESISTANCE TO EMB).
 FT VARIANT 959 959 D -> A (RESISTANCE TO EMB).
 FT VARIANT 1000 1000 M -> R (RESISTANCE TO EMB).
 FT VARIANT 1024 1024 D -> N (RESISTANCE TO EMB).
 FT VARIANT 773 774 SW -> FL (IN REF. 1).
 SQ SEQUENCE 1098 AA; 118020 MW; DD7D/025DC803833 CRC64;

Query Match: 8.9%; Score 81; DB 1; Length 1098;
 Best Local Similarity: 20.6%; Pred. No. 7.8; Mismatches: 50; Indels: 14; Gaps: 3;

QY 6 FLAWLPCMVPCMLPWRPWMSSTASWSSAALERSTOP-----ATGATWTKWAH 58
 Db 628 FLALAEFLALCAGATTGWWVVSYGVFNSAMPKIDITVSIFFALEIAANGYAALH 687
 QY 59 YACSSRSIPTLELTIVSR-----FLASLRVRYVCLRLCPYPKDS 101
 Db 688 FAPRGAGEGLRLLATTAAPVPIAGMAMAVFVSMVAGIV-ROVPTYSN 735

RESULT 4
 BA13_HUMAN STANDARD; PRM; 1522 AA.
 ID BA13_HUMAN STANDARD; PRM; 1522 AA.
 AC 060242; 060297; DT 16-OCT-2001 (Rel. 40, Created)
 AC 060242; 060297; DT 16-OCT-2001 (Rel. 40, Last sequence update)
 AC 060242; 060297; DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Brain-specific angiogenesis inhibitor 3 precursor.
 GN BA13 OR KIA0550.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC TISSUE=retal brain; RX MEDLINE:98194217; PubMed:9533023;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 CC -1- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION AND SUPPRESSION OF GLIOBLASTOMA.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOMORFS: A LONG FORM (SHOWN HERE) AND A SHORT FORM, MAY BE PRODUCED BY ALTERNATIVE SPlicing.
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN HEART. REDUCED EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL LINES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 CC -1- SIMILARITY: CONTAINS 4 TSP 1 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL: AB00529; BAA25363.1; -
 DR EMBL: AB01112; BAA25476.1; -
 DR Genew; HNC:945; BA13.
 DR MIM: 602684; -
 DR InterPro: IPRO000859; CUB_domain.
 DR InterPro: IPRO000832; GPCR_secretin.
 DR InterPro: IPRO000832; GPCR.
 DR InterPro: IPRO00203; PKC_cys-rich.
 DR InterPro: IPRO00884; TSP1.
 DR InterPro: IPRO01879; homm_receptor.
 DR Pfam: PF00002; 7tm_2; 1.
 DR Pfam: PF01825; GPS; 1.
 DR Pfam: PF01825; GPS; 1.
 DR InterPro: IPRO00203; PKC_cys-rich.
 DR InterPro: IPRO00884; TSP1.
 DR SMARF; SM00083; HOMM; 1.
 DR SMARF; SM00209; TSP1; 4.
 DR PROSITE: PS01180; CUB; 1.
 DR PROSITE; PS05021; GPS; 1.
 DR PROSITE; PS00649; G_Protein_Recep_F2_1; FALSE_NEG.

RESULT 7		ID	M2A2_HUMAN	STANDARD;	PRT;	1139 AA.
EMB_MCAV	STANDARD;	PRT;	1065 AA.			
EMB_MCAV	STANDARD;	PRT;	1065 AA.			
TD		AC	49641; Q17954;			
P71486;		DT	01-FEB-1996 (Rel. 33, Created)			
15-JUN-2002 (Rel. 41, Created)		DT	15-JUL-1998 (Rel. 36, Last sequence update)			
15-JUN-2002 (Rel. 41, Last sequence update)		DT	15-JUN-2002 (Rel. 41, Last annotation update)			
15-JUN-2002 (Rel. 41, Last annotation update)		DE	Alpha-mannosidase IIX (EC 3.2.1.114) (Mannosyl-oligosaccharide 1,3-			
DE		DE	1,6-alpha-mannosidase) (MAN IIX) (Mannosidase alpha class 2A member			
DE		DE	2).			
DE		GN	MANA2 OR MANA2X.			
OS		OS	Homo sapiens (Human).			
OC		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;			
OC		OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX		OX	NCBI_TaxID=9606;			
RN		[1]	SEQUENCE FROM N.A.			
RP		[1]	SEQUENCE FROM N.A. (LONG AND SHORT FORMS).			
RC		RP	TISSUE=Melanoma;			
RC		RC	Medline=96102195; PubMed=8524845;			
RA		RA	Medline=96102195; PubMed=8524845;			
RA		RA	MIAGO M., Liao Y.-F., Kudo S., Eto S., Mattei M.-G., Moremen K.W.,			
RA		RA	Futuda M.N.;			
Brennan P.J., Inamine J.M.;		RT	"Molecular cloning and expression of cDNAs encoding human alpha-mannosidase II and a previously unrecognized alpha-mannosidase IIX			
The embAB genes of Mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol.";		RT	isozyme.";			
Belanger A.E., Basra G.S., Ford M.E., Mikusova K., Belisle J.T., Brenner J.M.;		RT	Proc. Natl. Acad. Sci. U.S.A. 93:11919-11924(1996).			
-I- FUNCTION: Arabinosyl transferase responsible for the polymerization of arabinose into the arabinan of arabinogalactan.		CC	-I- POLYMERIZATION: Arabinosyl transferase (Probable).			
-I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).		CC	-I- SIMILARITY: BELONGS TO THE EMB FAMILY.			
CC		CC	CC			
CC		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC		CC	CC			
CC		CC	DR; U66560; AAC44568.1; -			
KW		KW	transf erase; Glycosyltransferase; Transmembrane; Cell wall; antibiotic resistance.			
FT		FT	POTENTIAL.			
TRANSMEM	15	TRANSMEM	15	POTENTIAL.		
FT		FT	204	226	POTENTIAL.	
TRANSMEM	241	TRANSMEM	241	263	POTENTIAL.	
FT		FT	394	413	POTENTIAL.	
TRANSMEM	411	TRANSMEM	411	436	POTENTIAL.	
FT		FT	510	527	POTENTIAL.	
TRANSMEM	510	TRANSMEM	510	557	POTENTIAL.	
FT		FT	567	589	POTENTIAL.	
TRANSMEM	596	TRANSMEM	596	618	POTENTIAL.	
FT		FT	633	655	POTENTIAL.	
TRANSMEM	667	TRANSMEM	667	689	POTENTIAL.	
FT		FT	1065 AA; 114624 MW; 3F12D13678C62BF CRC64; SEQUENCE			
SQ		Query Match	8.6%; Score 78.5; DB 1; Length 1065;			
Best Local Similarity	23.8%	Pred. No. 13;	Pred. No. 13; Mismatches 13; Conservative Matches 30;			
Matches		Indels	46; Indels 37; Gaps 4;			
QY	6	FLAWLPCMVNCWLPRWTWWNSSSSTAWNSWASSALESTSTOPATG-----	50			
Dy	597	FLAALPFMMLCAFATTGWWTVSS-----YGFPNSTMPKIGGITVVFMSVFAA	648			
QY	51	ATWKPLHYAGSSRSRISPTLETLTVSP-----FLASLRVARCURLCPPYPKDSSTE	104			
Dy	649	ALYALMHLFASEHGRGLRALTAAPVPLAGFMALVFTASVMAGIV-----ROYP	700			
QY	105	SWRVAW 110				
Dy	701	TYSNW 706				
QY	21	WR-TW---WWSSTAWNSWASSALESTSTOPATGATWT-----	55			
Query Match	8.6%; Score 78.5; DB 1; Length 1139;					
Best Local Similarity	21.1%	Pred. No. 13;	Pred. No. 13; Mismatches 35; Conservative Matches 35;			
Matches		Indels	53; Gaps 7;			

QY 38 SSALETSTQP---ATGATWTKWILHYAGSSRISPTPLEAHL-TWAPFLASLRVARVCLRLIC 93
 QY ||| : | | | :: | | | : | | | : | | | : | | |
 Db 1578 SSRQFSMMSPHYSSETSPSYSPTSPPAAGSPVPSYSPPSPTSPSPTSPSPTSPS 1637
 QY 94 PPY ---PKDSSTERPSWRVAWPSCPASPASPAQLMSSPRWMP 130
 QY ||| : | | | : | | | : | | | : | | | : | | |
 Db 1638 PSVSPTSPSYSPSPSYSPSSPSPSYSPSSPRYSPT 1678
 QY ||| : | | | : | | | : | | | : | | | : | | |
RESULT 13
CYSZ-SALTY STANDARD; PRT: 253 AA.
 AC P12673;
 DT 01-OCT-1989 (Rel. 12, created)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE CYSZ protein.
 GN CYSZ OR STM2429.
 OS *Salmonella typhimurium*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 RT *Salmonella*.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL STRAIN=LJ2;
 RX MEDLINE=88257033; PubMed=3209198;
 RA Byrne C.R., Monroe R.S., Ward K.A., Kredich N.M.,
 RT "DNA sequences of the cysK regions of *Salmonella typhimurium* and
 Escherichia coli and linkage of the cysK regions to ptsH.";
 RL J. Bacteriol. 170:3150-3157(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LJ2 / SGSC412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11577609;
 RA McClelland M., Sanderson K.E., Spelth J., Clifton S.W., Latreille P.,
 Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 Leonard S., Nguyen C., Scott K., Holmes A., Grawell N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: POSSIBLY INVOLVED IN SULFATE TRANSPORT.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -1- SIMILARITY: BELONGS TO THE CYSZ FAMILY.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 European Bioinformatics Institute. There are no restrictions on its
 use by non profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 EMBL: M21450; AAA27050.1; ALT_INIT.
 EMBL: AEO0880; AAI21323.1; -.
 DR A28181; BVERCZ.
 STYGENE: SG10081; cysz.
 DR KW Cysteine biosynthesis; Transport; Inner membrane; Transport;
 Complete proteome.
 FT TRANSMEM 31 POTENTIAL.
 FT TRANSMEM 72 POTENTIAL.
 FT TRANSMEM 151 171 POTENTIAL.
 FT TRANSMEM 222 242 POTENTIAL.
 SQ SEQUENCE 253 AA; 28906 MW; 101B2E2D48CC285 CRC64;
 QY Query Match 8.4%; Score 76.5; DB 1; Length 253;
 QY Best Local Similarity 22.2%; Pred No. 4.6; Matches 42; Conservative 28; Mismatches 64; Indels 55; Gaps 9;
 QY 8 AWLPCMM--VPCWLPWRTW-----WWSSSTAWWS-----WASSAL 41
 QY ||| : | | | : | | | : | | | : | | | : | | |
 Db 56 AWIPLSLMSHPDWLQLWLSYLWPAIVSLLVGFYFFSTLANNIAAPFGNLLABOLEARL 115
 QY 42 ETSQPATG3-----ATWTKWILHYAGSSRISPTL---EATITVSPLASLK 84
 QY ||| : | | | : | | | : | | | : | | | : | | |
 Db 115 TGATPPDTGIGLGMKDVRIMKREWKWLWLPRAIVLILYFPGIGTIAPWLFES 175
 QY 85 ARVCRLLCPYPKDSSTERPSWRVAWPSCPASPASLAQLMSSPRWMP---TCIUVTKML 139
 QY ||| : | | | : | | | : | | | : | | | : | | |
 Db 176 AWMLAIQYC-DYPFDNH---KVPFKTMRAALTQKVANMQGALTSLFTMIPVNLFI 229
 QY 140 RPWHAACGA 148
 QY ||| : | | |
 Db 230 MP-YAVCGA 237
RESULT 14
WFS1_HUMAN STANDARD; PRT: 890 AA.
 ID WFS1_HUMAN
 AC 076024;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 OX NCBI_TaxID=9606;
 RN Wolfflamin.
 RP WFS1.
 RC Homo sapiens (Human).
 RX Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RT [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99036570; PubMed=9817917;
 RA Strom T.M., Hoertnagel K., Hofmann S., Gekeler F., Scharfe C.,
 RA Rabl W., Gerbitz K.-D., Meitinger T.;
 RT "Diabetes insipidus, diabetes mellitus, optic atrophy and deafness
 (DIMODAD) caused by mutations in a novel gene (wolframin) coding for
 a predicted transmembrane protein";
 RT Hum. Mol. Genet. 7:2021-2028(1998).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS WFS LEU-504; VAL-695 AND LEU-724.
 RC TISSUE=Brain;
 RX MEDLINE=9842659; PubMed=9771706;
 RA Inoue H., Tanizawa Y., Wasson J.J., Behn P., Kalidas K.,
 RA Berni-Mizrachi E., Mueckle M., Marshall H., Donis-Keller H.,
 RA Crock P., Rogers D., Mikuni M., Kumashiro H., Higashi K., Sobue G.,
 RA Oka Y., Permutt M.A.;
 RT "A gene encoding a transmembrane protein is mutated in patients with
 diabetes mellitus and optic atrophy (Wolfran Syndrome).";
 RT diabetics mellitus and optic atrophy (Wolfran Syndrome).";
 RL Nat. Genet. 20:143-148(1998).
 RN [3]
 RP VARIANTS WFS, AND VARIANTS VAL-333 AND HIS-611.
 RX MEDLINE=99452889; PubMed=10521293;
 RA Hardy C., Khanim F., Torree R., Scott-Brown M., Seller A., Poulton J.,
 RA Collier D., Kirk J., Polymeropoulos M., Latif F., Barrett T.,
 RT "Clinical and molecular genetic analysis of 19 Wolfran syndrome
 RT kindreds demonstrating a wide spectrum of mutations in WFS1.>";
 RL Ann. J. Hum. Genet. 65:1279-1290(1999).
 RN [4]
 RP VARIANTS THR-59 AND HIS-611.
 RX MEDLINE=20088163; PubMed=1064825;
 RA Furkong R.A., Ho L.W., Rubinstein J.S., Michael A., Walsh C.,
 RA Paykel E.S., Rubinstein D.C.,
 RT "A rare coding variant within the Wolframin gene in bipolar and
 RT unipolar affective disorder cases.>";
 RL Neurosci. Lett. 277:123-126(1999).
 RN [5]
 RP VARIANTS HIS-456; SER-576; HIS-611 AND VAL-720.
 RX MEDLINE=2014581; PubMed=1067925;
 RA Awata T., Inoue K., Kurihara S., Ohkubo T., Inoue I., Abe T.,
 RA Takino H., Kanazawa Y., Katayama S.,
 RT "Missense variations of the gene responsible for Wolfran syndrome
 (WFS1/Wolframin) in Japanese: possible contribution of the Arg56His
 mutation to type 1 diabetes as a nonautoimmune genetic basis.>";

RL	Blochem. Biophys. Res. Commun.	268:612-616(2000).	FT	VARIANT	169	169	/FTId=VAR_011306.
RN	[6]		FT	VARIANT			E -> K (IN WFS).
RP	VARIANTS HIS-456; SER-576; HIS-611; VAL-720 AND LYS-737.		FT	VARIANT			/FTId=VAR_009109.
RX	MEDLINE-2025808; PubMed-076054;		FT	VARIANT	292	292	P -> S (IN WFS).
RA	Ohtsuki T., Ishiguro H., Yoshikawa T., Arinami T.;		FT	VARIANT			/FTId=VAR_009110.
RT	"WFS1 gene mutation search in depressive patients: detection of five missense polymorphisms but no association with depression or bipolar affective disorder.";		FT	VARIANT	296	296	I -> S (IN WFS).
RT	J. Affect. Disord. 58:11-17 (2000).		FT	VARIANT			/FTId=VAR_009111.
RL	VARIANT WFS ILE-443, AND VARIANTS VAL-333; HIS-611; VAL-684 AND CYS-708.		FT	VARIANT	333	333	/FTId=VAR_009112.
RX	MEDLINE-21193172; PubMed-11295031;		FT	VARIANT	333	333	I -> V.
RA	Tessa A., Carbone I., Matteoli C., Bruno C., Patrono C., Patera I.P., De Luca F., Lorini R., Santorelli F.M.;		FT	VARIANT	350	350	/FTId=VAR_005840.
RT	"Identification of novel WFS1 mutations in Italian children with Wolfram syndrome.";		FT	VARIANT	437	437	MISSING (IN WFS).
RT	Hum. Mutat. 17:348-349(2001).		FT	VARIANT	354	354	/FTId=VAR_011307.
RN	[8]		FT	VARIANT	415	415	MISSING (IN WFS).
RP	VARIANTS WFS VAL-58; THR-126; 350-PHE DEL; 354-PHE DEL; LEU-504; ARG-780 AND CYS-811, AND VARIANTS ARG-674 AND LYS-737.		FT	VARIANT	456	456	/FTId=VAR_009113.
RX	MEDLINE-21111358; PubMed-11161832;		FT	VARIANT	504	504	G -> R (IN WFS).
RA	Gomez-Zaera M., Strom T.M., Rodriguez B., Estivill X., Meitinger T., Nunes V.;		FT	VARIANT	504	504	/FTId=VAR_009114.
RA	"Presence of a major WFS1 mutation in Spanish Wolfram syndrome pedigree";		FT	VARIANT	443	443	S -> T (IN WFS).
RT	Mol. Genet. Metab. 72:7-81(2001);		FT	VARIANT			/FTId=VAR_011308.
RL	RT - SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum.		FT	VARIANT			R -> H.
CC			FT	VARIANT	567	568	/FTId=VAR_009115.
CC	- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART FOLLOWED BY BRAIN, PLACENTA, LONG AND PANCREAS. WEAKLY EXPRESSED IN ISLET AND BETA-CELL, INSULINOMA CELL LINE.		FT	VARIANT	576	576	G -> S, /FTId=VAR_005842.
CC	- POLYMORPHISM: R456H, R611H, AND I720V POLYMORPHISMS ARE IN TIGHT LINKAGE DISQUILIBRIUM WITH ONE ANOTHER AND ASSOCIATED WITH TYPE 1 DIABETES IN JAPANESE.		FT	VARIANT			A -> T (RARE; COULD BE A RISK FACTOR FOR AFFECTIVE DISORDER).
CC	(ALSO KNOWN AS DIABETES INSIPIDUS AND MELLITUS WITH OPTIC ATROPHY AND DEAFNESS SYNDROME (DIDMOAD). IT IS A RARE AUTOSOMAL RECESSIVE DISORDER CHARACTERIZED BY JUVENILE DIABETES MELLITUS, DIABETES INSIPIDUS, OPTIC ATROPHY, DEAFNESS AND VARIOUS NEUROLOGICAL SYMPTOMS.		FT	VARIANT			/FTId=VAR_010603.
CC			FT	VARIANT	611	611	R -> H.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		FT	VARIANT			/FTId=VAR_005843.
CC			FT	VARIANT	674	674	MISSING (IN WFS).
CC			FT	VARIANT	674	674	/FTId=VAR_009115.
CC			FT	VARIANT	684	684	/FTId=VAR_005842.
CC			FT	VARIANT	684	684	/FTId=VAR_011309.
CC			FT	VARIANT	690	690	A -> V.
CC			FT	VARIANT	695	695	G -> V (IN WFS).
CC			FT	VARIANT	695	695	/FTId=VAR_005844.
CC			FT	VARIANT	700	700	W -> C (IN WFS).
CC			FT	VARIANT	700	700	/FTId=VAR_009117.
CC			FT	VARIANT	708	708	R -> C.
CC			FT	VARIANT	720	720	C -> R (IN WFS).
CC			FT	VARIANT	724	724	/FTId=VAR_009116.
CC			FT	VARIANT	724	724	G -> V (IN WFS).
CC			FT	VARIANT	736	736	P -> L (IN WFS).
CC			FT	VARIANT	736	736	/FTId=VAR_010604.
CC			FT	VARIANT	737	737	P -> L (IN WFS).
CC			FT	VARIANT	737	737	/FTId=VAR_011312.
CC			FT	VARIANT	780	780	G -> R (IN WFS).
CC			FT	VARIANT	818	818	/FTId=VAR_011313.
CC			FT	VARIANT	818	818	R -> C (IN WFS).
CC	KW	Transmembrane; Polymorphism; Disease mutation; Diabetes mellitus; Diabetics insipidus; Deafness.	FT	VARIANT			/FTId=VAR_011314.
FT	POTENTIAL.		FT	VARIANT	885	885	P -> L (IN WFS; MILD FORM).
FT	POTENTIAL.		FT	SEQUENCE	890	AA:	/FTId=VAR_009119.
FT	POTENTIAL.		FT	SEQUENCE	890	AA:	1 TWRFLFLAWLPCMMMPCW-----LWRTWWWSSTAW-VSWASSALE 43
FT	POTENTIAL.		FT	SEQUENCE	890	AA:	Db 631 SMVKLLWVIAVILPCWFYVRSGRMKVNSTLWQYQALCGPRAWKETNMARTQILC 690
FT	POTENTIAL.		FT	SEQUENCE	890	AA:	OY 44 STQPATGATWKLHYAGSSRSIPLATEATVSPF 78
FT	POLY-LYS.		FT	SEQUENCE	890	AA:	Db 691 SHLEGHRVWTWGRFKVVRVDIDNSAESANMLPF 725
FT	POLY-PHE.		FT	SEQUENCE	890	AA:	A -> V (IN WFS).
FT	VARIANT	126	126			A -> T (IN WFS).	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

Om protein - protein search, using sw model

Run on:

March 27, 2003, 16:01:02 ; Search time 47.3684 Seconds

(without alignments)
704.681 Million cell updates/sec

Title: US-10-019-219-1

Perfect score: 912

Sequence: TVVRFLFLAWLPCMMVPCWLPA.....WAACCGARVKRRFLQLTLSR 162

Scoring table: BLOSUM62

gapop 10.0 , Gapext: 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%, Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21:*

1: sp_archea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp Rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp unclassified:*

15: sp_virus:*

16: sp_bacterian:*

17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Length DB ID Description

ALIGNMENTS

RESULT 1	Q9MIA8	PRELIMINARY,	PRT;	343 AA.
ID	Q9MIA8;			
AC				
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)		
DT	01-OCT-2000	(TREMBLrel. 21, Last annotation update)		
DE	NADH dehydrogenase subunit 2.			
GN	NADH2.			
OS	Typhlonectes natans (Rubber eel).			
OG	Mitochondrion.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Gymnophiona; Caeciliidae; Typhlonectidae.			
OX	NCBL-TaxID=8456;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=2029652; PubMed=10835397;			
RA	Zardoya R., Meyer A.;			
RT	"Mitochondrial Evidence on the Phylogenetic Position of Caecilians (Amphibia: Gymnophiona);"			
RT	(Amphibia: Gymnophiona);"			
RL	Genetics 155:765-775(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Zardoya R.;			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.			
CC	- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+)			
.DR	EMBL: AF154051; AAC78184.1; -.			
DR	InterPro: IPR01750; Oxidored_q1.			
DR	Protein: PFO0361; oxidored_q1; 1.			
KW	Mitochondrion; NAD; Oxidoreductase; Ubiquinone.			
SQ	SEQUENCE 343 AA; 37519 MW; 2BAD5B15406BB2C8 CRC64;			
	Query Match Best Local Similarity 10.7%; Score 98; DB 8; Length 343; Matches 32; Conservative 13; Mismatches 56; Indels 22; Gaps 4;			
QY	SSSTAAWSASSA-----LETSTOPATGATWKW-LHYAGSSRSRSPTEATLTVSPF 78			
DB	Q8WTF5 crotalus mi			

ID	Q95QF5	PRELIMINARY;	PRT;	842 AA..
AC	095QF5;			
DR	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DE	Hypothetical 95.0 kDa protein.			
GN	F49E10.2.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nemataoda; Chromadorea; Rhabditida; Rhabditoidea;			
RC	Rhabditidae; Peloderrinae; Caenorhabditis;			
RA	NCBL-TaxID=6239;			
RT	"None;"			
SEQUENCE FROM N.A.				
RL	STRAIN=BRISTOL N2;			
RC	STRAIN=99069613; Pubmed=9851916;			
RA	None;			
RT	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";			
RL	Science 282:2012-2018(1998).			
RN	[12]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BRISTOL N2;			
RA	Waterson R.;			
RT	"Direct Submission;"			
RL	Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; 05341; AAL06020.1; -.			
DR	InterPro; IPR04294; RP065.			
DR	Pfam; PF03055; RPE65; 1.			
KW	Hypothetical protein.			
SQ	SEQUENCE 842 AA; 94955 MW; 3B0230FB7DB110CA CRC64;			
RESULT	9			
Q95QF5	Best Local Similarity 24.8%; Pred. No. 4.8; Mismatches 39; Indels 39; Gaps 4;			
AC	Q95QF5;	PRELIMINARY;	PRT;	195 AA.
DR	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DE	Alpha-amylase.			
GN	AMY-D OR CG17076.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Arthropoda; Insecta; Drosophilidae; Drosophila.			
OC	Subarygota; Metazoa; Arthropoda; tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriodea; Drosophilidae; Drosophila.			
OX	NCBL-TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=KN7;			
RA	MEDLINE=2110343; Pubmed=11156987;			
RT	Aroki H.; Inomata N.; Yamazaki T.;			
RT	"Molecular evolution of duplicated Amylase gene regions in Drosophila melanogaster: Evidence of positive selection in the coding regions and selective constraints in the cis-regulatory regions.;"			
RL	Genetics 157:667-677(2001).			
DR	EMBL; ABO43038; BAB32536.1; -.			
DR	HSSP; P26634; IJAE;			
DR	FLYbase; FBgn000078; Amy-d.			
SQ	SEQUENCE 195 AA; 20734 MW; OCBOFE3660837C20 CRC64;			
RESULT	9			
Q95QF5	Best Local Similarity 26.0%; Pred. No. 1.3; Mismatches 34; Conservative 18; Indels 42; Gaps 37;			
AC	Q95QF5;	PRELIMINARY;	PRT;	890 AA.
DR	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
GN	WFS1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OC	NCBL-TaxID=10116;			
RC	SEQUENCE FROM N.A.			
RT	TESSUE=Pancreas;			
RX	MEDLINE=21096923; PubMed=11181571;			
RA	Takeda K., Inoue H., Tanizawa Y., Matsuzaki Y., Oba J., Watanabe Y., Shinoda K., Oka Y.;			
RT	"WIF1 (Wolfram syndrome 1) gene product: predominant subcellular localization to endoplasmic reticulum in cultured cells and neuronal expression in rat brain;"			
RT	Hum. Mol. Genet. 10:477-484(2001).			
RL				
RESULT	11			
Q95QF5	Best Local Similarity 21.2%; Pred. No. 5.1; Mismatches 45; Indels 39; Gaps 5;			
AC	Q95QF5;	PRELIMINARY;	PRT;	230 AA.
DR	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DE	InterPro; IPR02088; PPTA.			
GN	Prosite; PS00904; PPTA; UNKNOWN_1.			
OS	Euksyota; Metazoa; Nemataoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderrinae; Caenorhabditis;			
RC	NCBL-TaxID=6239;			
RA	None;			
RT	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";			
SEQUENCE FROM N.A.				
RL	STRAIN=BRISTOL N2;			
RC	STRAIN=99069613; Pubmed=9851916;			
RA	None;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[12]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[13]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[14]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[15]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[16]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[17]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[18]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[19]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[20]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[21]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[22]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[23]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[24]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[25]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[26]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[27]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[28]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[29]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[30]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[31]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[32]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[33]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[34]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[35]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[36]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[37]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[38]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[39]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[40]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[41]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[42]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[43]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[44]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[45]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[46]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[47]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RA	Miller N.;			
RT	"The sequence of <i>C. elegans</i> cosmid F49E10.;"			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[48]</td			

DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RA Strausberg R.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC012899; AAK12899.1; -;

KW HYPOTHETICAL PROTEIN.

SEQUENCE 230 AA; 23781 MW; B8F2194AC35DA358 CRC64;

Query Match 9.5%; Score 86.5; DB 4; Length 230; Best Local Similarity 23.6%; Pred. No. 1.6; Mismatches 34; Conservative 11; Indels 39; Gaps 6; Matches 34; NTER 1; CTER 1; Non_TER 1; SEQUENCE 270 AA; 30295 MW; 5736E887B6D12E5F CRC64;

QY 9 WLPCKMMVPCWLWPWTWWMSSSRAWNSWASSALETSSTOPATGATWKWLHYAGSSRISPT 68

Db 68 WVPATACSCC--PQAACWPPPTPAWT--AESVLAARSAGPA----- 103

QY 69 LPLTLITVSPLFLASLRVARVCLRLICPPYPKDSSTEPEHSPHRVANFSCPSPSLPAQIMSSPRWW 128

Db 104 ---LTPSPSTANTTAAPGAVERSCPQAQSASWRMTVGRAASST---AMPGAHWA 155

QY 129 PTCGLPVTKLRLD----WWAAC 146

Db 156 AVQRR-SALTRAPPSCPQSAWLPC 178

RESULT 12

ID 080FX1 PRELIMINARY; PRT; 1325 AA.

AC 080FX1; DE 01-JUN-2002 (TREMBrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBrel. 21, Last sequence update)

DE RIM-binding protein 2.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OX NCBI_TAXID=9031;

RN [1]

SEQUENCE FROM N.A.

RA Hibino H., Pironkova R., Onumere O., Vologdskaia M., Hudspeth A.J., Lesage F.;

RT RIM-binding Proteins (RBPs) couple Rab3-interacting molecules (RIMs) to voltage-gated Ca²⁺ channels.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EML1; AV072908; AAL6795.1; -;

SQ SEQUENCE 1325 AA; 147537 MW; 1DD8C50D4CDA1B5B CRC64;

Query Match 9.4%; Score 86; DB 13; Length 1325; Best Local Similarity 30.4%; Pred. No. 11; Mismatches 35; Conservative 14; Indels 18; Gaps 4; Matches 35; NTER 1; CTER 1; Non_TER 1; SEQUENCE 210 AA; 22143 MW; 83FF8346397EEF043 CRC64;

QY 27 SSSSTAWWSSAALESTQPATGATWKWLHYAGSSRIS---PLETLITYVSPLFLASL 82

Db 517 STQATQVSKWPKLICPPYK-DSSTEPSWRVAVPSCPSPSLPAQIMSSPRWWTCPLPYWK 136

QY 83 RVARVCLRLICPPYK-DSSTEPSWRVAVPSCPSPSLPAQIMSSPRWWTCPLPYWK 136

Db 576 EAKEVTVPLSLQAGESVDSSV-----AAIPSDLLVPPSPHPRTAPSK 618

RESULT 13

ID 0917M0 PRELIMINARY; PRT; 270 AA.

AC 0917M0; DE 01-MAR-2001 (TREMBrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBrel. 16, Last sequence update)

DE FL000050 protein (Fragment).

GN FL000050.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=SPLLEN;

RA Ohara O., Nakase T., Kikuno R., Okumura K.; DE "The nucleotide sequence of a long cDNA clone isolated from human spleen.";

RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AR024458; BAB15748.1; -;

SQ SEQUENCE 270 AA; 30295 MW; 5736E887B6D12E5F CRC64;

Query Match 9.4%; Score 85.5; DB 4; Length 270; Best Local Similarity 23.0%; Pred. No. 2.3; Mismatches 42; Conservative 12; Indels 35; Gaps 9; Matches 35; NTER 1; CTER 1; Non_TER 1; SEQUENCE 270 AA; 30295 MW; 5736E887B6D12E5F CRC64;

QY 19 LPWTIWW-WSSSTAWWSSAALESTQPATGATWKWLHYAGSSRISPT 68

Db 98 LSMSITWEEQERLSQSWALEDNLSVLRNQTSKMEELKKWDLRSPGELERGSSRVSP-- 154

QY 71 ATLTVSPFLASLRVA---RVCLRLICPPYKSS---TEPS-----WRVAMPSCPASTL 117

Db 155 --GIRPLCEGDRIALPSPSPAQLOGIMPSSAPQOKSLVLFQPGTGLSLWRPRPSSKG 211

QY 118 P-----AQLMSSPR--WMPTCLV 134

Db 212 PQNPGQQQALWVAQRIKWWPPCCFV 236

RESULT 14

ID 0919F6 PRELIMINARY; PRT; 210 AA.

AC 0919F6; DE 01-NOV-1999 (TREMBrel. 12, Last sequence update)

DT 01-NOV-1999 (TREMBrel. 12, Last sequence update)

DE 01-MAR-2002 (TREMBrel. 20, Last annotation update)

QY 01-AUG-2002 (TREMBrel. 20, Last annotation update)

DE HYPOTHETICAL PROTEIN ARF2332.

GN APE2332.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales; Desulfurococcaceae; Aeropyrum.

OX NCBI_TaxID=56336;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=KL;

DR MEDLINE-99310339; PubMed=10382966;

SQ SEQUENCE 1325 AA; 147537 MW; 1DD8C50D4CDA1B5B CRC64;

Query Match 9.4%; Score 86; DB 13; Length 1325; Best Local Similarity 30.4%; Pred. No. 11; Mismatches 35; Conservative 14; Indels 18; Gaps 4; Matches 35; NTER 1; CTER 1; Non_TER 1; SEQUENCE 210 AA; 22143 MW; 83FF8346397EEF043 CRC64;

QY 27 SSSSTAWWSSAALESTQPATGATWKWLHYAGSSRIS---PLETLITYVSPLFLASL 82

Db 517 STQATQVSKWPKLICPPYK-DSSTEPSWRVAVPSCPSPSLPAQIMSSPRWWTCPLPYWK 136

QY 83 RVARVCLRLICPPYK-DSSTEPSWRVAVPSCPSPSLPAQIMSSPRWWTCPLPYWK 136

Db 576 EAKEVTVPLSLQAGESVDSSV-----AAIPSDLLVPPSPHPRTAPSK 618

Query Match 9.3%; Score 84.5; DB 17; Length 210; Best Local Similarity 25.0%; Pred. No. 2.2; Mismatches 34; Conservative 26; Indels 19; Gaps 6; Matches 34; NTER 1; CTER 1; Non_TER 1; SEQUENCE 210 AA; 22143 MW; 83FF8346397EEF043 CRC64;

QY 25 WSSSTAWWSSAALESTQPATGATWKWLHYAGSSRISPLETLITYVSPLFLASL 84

Db 74 WSTASST-TSFKSSVDDFTSGRTMSSTPSF---SASFLLAASSAPLTTMGPIPEIRL 127

Db	128	FYIACLTALISAPAFSTSEYLEPPFSATMSISPTSPSSMGT-----WKS-----RNLK	178
Oy	141	PWWAACGARVKRHFLQ 156	
Db	179	PDFSISTRAAISSIFLQ 194	
			RESULT 15
OYXK16		: : : : : : : : : : :	85 ARVCLRLCPYPKPYKDSST---EPSWRVIAWPSCPASLDQLMSSPRWWPPCLPVTKLIR
ID 09KX16		PRELIMINARY;	551 AA.
AC O9KX16;			
DT 01-OCT-2000	(TREMBREL)	15' Created)	
DR 01-JUN-2002	(TREMBREL)	15' Last sequence update)	
DE Putative integral membrane protein.			
GN SCO2334 OR SCC53_25.			
OS Streptomyces coelicolor.			
OC Bacteria; Firmicutes; Actinobacteria; Streptomycetaceae; Streptomyces.			
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.			
OX NCBI_TAXID=1902;			
[1]			
RP SEQUENCE FROM N.A.			
RC STRAINA3(2) / M145;			
RA Bentley S.D., Chatter K.F., Cerdeno-Tarraga A.-M., Challis G.L.,			
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,			
RA Harper D., Battemann A., Brown S., Chandra G., Chen C.W., Collins M.,			
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,			
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,			
RA Rabbowitsch E., Ralanderam M.A., Rutherford K., Rutten S.,			
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,			
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,			
RA Hopwood D.A.;			
RT "Complete genome sequence of the model actinomycete Streptomyces			
coelicolor A3(2)",			
Nature 417:141-147(2002);			
RL EmbL: AL357591; CAB9454.1; -.			
DR EMBL: AL357591; CAB9454.1; -.			
SQ SEQUENCE 551 AA; 59486 MW; 6DF4A5BE23D8E9E CRC64;			
Query Match 9.3%; Score 84.5; DB 16; Length 551;			
Best Local Similarity 24.7%; Pred. No 5.9;			
Matches 44; Conservative 15; Mismatches 46; Indels 73; Gaps			
Oy 7 LAWLPCMVPCPLWPWTWWSST-----ANYSWASSALETSTQPANGATWIKWL 57			
: : : : : : : : : : : : :			
35 LATLPT----IPLYVWMAFLATGGGLAQEAEDFAS---RHGGSGAYGLFWYGMG 84			
Oy 58 HYAGSSS1SPILER----TLV----SPFLASLRVARWCURLCPYPKPUSSTESWR 107			
: : : : : : : : : : : :			
85 HTANYSVISPLMLAVALGYRTVIVGGLAASWLAALVLYRVCGR-----R 1288			
Oy 108 VAWPSCPASLPQALMSSPRNN-----PTCILPVTKLRLPWWAAGC 147			
: : : : : : : : : : : :			
129 PWPWALLASLAI----WEDVASGRATAFLGAVALALAACYPLVR--ERRWLAAG 176			
Db			

Search completed: March 27, 2003, 16:12:05
Job time : 50.3684 secs

Job time : 50.3684 secs

GenCore version 5.1.4-p5-4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.4-p5-4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 15:54:52 ; Search time 5.42105 Seconds

(without alignments) (221.222 Million cell) updates/sec

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMWPCTCL 9

Scoring table: BIOSUM62

Gap pop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101022:*

1: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1980.DAT:*

2: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1981.DAT:*

3: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1982.DAT:*

4: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1983.DAT:*

5: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1984.DAT:*

6: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1985.DAT:*

7: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1986.DAT:*

8: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1987.DAT:*

9: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1988.DAT:*

10: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1989.DAT:*

11: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1991.DAT:*

12: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1992.DAT:*

13: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1993.DAT:*

14: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1994.DAT:*

15: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1995.DAT:*

16: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1996.DAT:*

17: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1997.DAT:*

18: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1998.DAT:*

19: /SIDS2/gcadata/geneseq/geneseq/emb1/AA1999.DAT:*

20: /SIDS2/gcadata/geneseq/geneseq/emb1/AA2000.DAT:*

21: /SIDS2/gcadata/geneseq/geneseq/emb1/AA2001.DAT:*

22: /SIDS2/gcadata/geneseq/geneseq/emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	63	100.0	9 22	AAB31702
2	63	100.0	162 22	AAB31701
3	63	100.0	156 22	AAB31703
4	46	73.0	72 22	AAB21874
5	46	73.0	105 19	AAW8196
6	45	73.0	105 19	RAW6156
7	45	71.4	102 22	RAW61607
8	45	71.4	107 22	AM06855
9	44	69.8	135 21	AAGC2709
10	43	68.3	49 20	AYZ25308

RESULT ID	ID	AA
1	AAB31702	standard; peptide; 9 AA.
2	XX	AC
3	XX	DT
4	30-APR-2001	(first entry)
5	XX	DE
6	XX	Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL; cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma; Eucalyptus grandis
7	XX	KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
8	OS Homo sapiens.	
9	XX	WO200100784-A2.
10	XX	PD 04-JAN-2001.
11	XX	PP 27-JUN-2000; 2000WO-FR01791.
12	XX	PR 28-JUN-1999; 99FR-000224.
13	XX	PA (INSR) INST ROUSSY GUSTAVE.
14	XX	PI Ronin C, Scott V, Triebel F;
15	XX	DR WPI; 2001-112443/12.
16	XX	PT New Peptides and its encoding nucleic acid derived from intestinal carboxyl esterase, useful as immunostimulants for treating cancer -
17	XX	PT HCV NS5B carboxy-t

11	43	68.3	237 22	AAB85763	Human protein seq
12	43	68.3	238 22	AAB61464	Human purified seq
13	43	68.3	817 19	AAW8101	Human discs large
14	43	68.3	817 20	AAV10137	Human acidic sequenc
15	43	68.3	849 19	AAW8102	Human discs large
16	43	68.3	3010 20	AAW98022	Infectious hepatitis
17	43	68.3	3010 22	AAW31170	Amino acid sequenc
18	43	68.3	3010 22	AAB59174	Protein encoded by
19	42	66.7	34 22	ABB43800	Peptide #11745 enc
20	42	66.7	34 22	ABB6738	Protein #11306 enc
21	42	66.7	34 22	AAB64766	Human brain express
22	42	66.7	34 22	AAM7529	Human bone marrow
23	42	66.7	34 22	AAM21459	Peptide #7893 enco
24	42	66.7	34 22	AAM37708	Propionibacterium
25	42	66.7	34 23	AAB6551	Human peptide enco
26	42	66.7	58 22	AAU22306	Propionibacterium
27	42	66.7	112 22	AAM24343	Novel human diago
28	42	66.7	124 22	ABB15805	Human ovarian anti
29	41	65.1	50 22	AAB62798	Propionibacterium
30	41	65.1	58 22	AAC61428	Human immune/haema
31	41	65.1	68 21	AAB56355	Propionibacterium
32	41	65.1	94 22	AAU22567	Propionibacterium
33	41	65.1	96 22	ABG08139	Human EST encoded
34	41	65.1	124 22	ABB15805	Human ORF protein
35	41	65.1	124 23	ABP05057	Human peptide
36	41	65.1	146 22	AMB8148	Propionibacterium
37	40	63.5	96 22	AAB50855	Propionibacterium
38	40	63.5	106 22	AAU22567	Propionibacterium
39	40	63.5	20 20	AYX88684	Human yb3-1 secre
40	40	63.5	155 22	AAU39040	Human secreted pro
41	40	63.5	155 23	ABB5749	Human polypeptide
42	40	63.5	158 22	AABM3974	Human immune/haema
43	40	63.5	293 21	ABR32804	Eucalyptus grandis
44	40	63.5	447 22	ABG19598	Novel human diagno
45	40	63.5	518 23	AAU93036	Arabidopsis trans

ALIGNMENTS

PS Claim 2; Page 33; 53p; French.

XX
CC The present sequence is derived from a human intestinal carboxylesterase
CC (ICE) polypeptide. ICE induces specific-specific cytotoxic T lymphocytes
(CTL) and secretion by these CTL of cytotoxic factor, e.g.
CC interleukin-2, interferon gamma and tumour necrosis factor. ICE
CC polypeptides and polynucleotides are used for treating cancer, by in
vivo or in vitro immunisation, particularly solid cancers and most
especially hepatocarcinoma or adenocarcinoma of colon and kidney. They
are also used to stimulate the immune system, and to increase, in
culture, the production of associated-associated CTL, for re-injection,
and/or to induce secretion of cytotoxic factors from CTL. Dendritic
CC cells loaded with ICE are used to induce such CTL in cultures.
XX

SQ Sequence 9 AA:
Query Match 100.0%; Score 63; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 0.026; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SPRWWPTCL 9
Db 124 SPRWWPTCL 132

RESULT 3
ID AAB31703 standard; Protein; 166 AA.
XX
AC AAB31703;
XX
DT 30-APR-2001 (first entry)
DE Protein encoded by an intestinal carboxylesterase (ICE) cDNA.
XX
KW Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;
KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
XX
OS Homo sapiens.
XX
PN WO200100784-A2.
XX
PD 04-JAN-2001.
XX
PP 27-JUN-2000; 2000WO-FR01791.
XX
PR 28-JUN-1999; 99FR-0008224.
XX
(INSR) INST ROUSSEY GUSTAVE.
PA Ronsin C, Scott V, Triebel F;
XX
DR WPI; 2001-112443/12.
DR N-PSDB; AAF25258.
XX
PT New peptides and its encoding nucleic acid derived from intestinal
PT carboxylesterase, useful as immunostimulants for treating cancer -
PS Disclosure; Fig 8A; 53p; French.

XX
CC The present sequence is encoded by the coding region of human intestinal
CC carboxylesterase (ICE) gene. ICE induces specific-specific cytotoxic T
CC lymphocytes (CTL) and secretion by these CTL of cytotoxic factors,
CC e.g. interleukin-2, interferon gamma and tumour necrosis factor. ICE
CC polypeptides and polynucleotides are used for treating cancer, by in
vivo or in vitro immunisation, particularly solid cancers and most
especially hepatocarcinoma or adenocarcinoma of colon and kidney. They
are also used to stimulate the immune system, and to increase, in
culture, the production of associated-associated CTL, for re-injection,
and/or to induce secretion of cytotoxic factors from CTL. Dendritic
CC cells loaded with ICE are used to induce such CTL in cultures.
XX

SQ Sequence 166 AA:
Query Match 100.0%; Score 63; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.027; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SPRWWPTCL 9
Db 128 SPRWWPTCL 136

RESULT 4
ID AAU21874 standard; Protein; 72 AA.
XX

AC AAU21874;
 XX
 DT 17-DEC-2001 (first entry)
 XX
 DE Human cardiovascular system antigen polypeptide SEQ ID No 648.
 XX
 KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
 chicken; sheep; immunosuppressive; antiarthritic; vasoactive; dog;
 antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
 cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
 ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
 hyperproliferative disorder; breast; liver; cardiovascular disorder;
 cerebrovascular disorder; nervous system disorder; bacterial infection;
 fungal infection; viral infection; ocular disorder; neoplasm;
 gastrentestinal disorder; renal disorder; respiratory disorder;
 wound healing; skin aging; organ transplantation; tissue regeneration;
 KW anti-infertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200155321-A2.
 XX
 PD 02-AUG-2001.
 XX
 PR 17-JAN-2001; 2001WO-US01340.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186330.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0220953.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0224513.
 PR 14-AUG-2000; 2000US-0224514.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0226219.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0233141.
 PR 08-SEP-2000; 2000US-0233144.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0233065.
 PR 14-SEP-2000; 2000US-0233298.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233498.
 PR 14-SEP-2000; 2000US-0233499.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0235484.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244674.
 PR 08-NOV-2000; 2000US-0244675.
 PR 08-NOV-2000; 2000US-0244676.
 PR 08-NOV-2000; 2000US-0244677.
 PR 08-NOV-2000; 2000US-0244678.
 PR 08-NOV-2000; 2000US-0245623.
 PR 08-NOV-2000; 2000US-0245624.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 08-NOV-2000; 2000US-0246617.
 PR 08-NOV-2000; 2000US-0246618.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.

XX Chen ISY, Jowett JBM, Withers-ward E;
 PT
 XX WPI: 1998-447375/38.
 XX
 PT Identification of compounds binding the HIV-1 Vpr protein - that
 PT block Vpr-mediated cell stasis, useful for treating HIV-1 or other
 PT lentiviral infections
 PS disclosure; Fig 8; 63pp; English.
 XX
 CC This represents the amino acid sequence of the B29-1 protein. This
 CC protein can bind to the Vpr protein encoded by the HIV genome. The
 CC invention provides a method of identifying an agent for use in treating
 CC lentiviral infections. The method comprises contacting a cellular target
 CC of the Vpr protein with the agent to be tested, and assessing the ability
 CC of the agent to block interaction of the Vpr protein with the cellular
 CC target, where an agent which blocks this interaction is an anti-
 CC lentiviral agent. Alternatively, the agent contacts a cell expressing the
 CC Vpr protein under conditions where the Vpr protein induces cell stasis in
 CC the absence of the agent. Identification of the agent is then observed
 CC by blockage of Vpr-induced cell stasis. The method allows the
 CC identification of compounds that block Vpr-mediated cell stasis and
 CC ultimately cell death. The compounds can thus be used in the treatment
 CC of HIV and other lentiviral infections.
 XX
 SQ Sequence 105 AA;
 Query Match 73.0%; Score 46; DB 19; Length 105;
 Best Local Similarity 75.0%; Pred. No. 6.6; Mismatches 6; Conservative 0; Indels 2; Gaps 0;
 Matches 6; Mismatches 0; Gaps 0;
 QY 1 SPRKWWPTC 8
 ID 60 SRRWWPAC 67
 XX
 RESULT 7
 AAM06407 ID AAM06407 standard; Protein; 102 AA.
 XX
 AC AAM06407;
 XX
 DT 05-OCT-2001 (first entry)
 DE Human foetal protein, SEQ ID NO: 138.
 XX
 KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 KW nootropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antisense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation.
 OS Homo sapiens.
 XX
 PN WO200155339-A2.
 XX
 PD 02-AUG-2001.
 XX
 PR 25-JAN-2001; 2001WO-US02723.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 15-SEP-2000; 2000US-0663870.
 PR 06-NOV-2000; 2000US-0707351.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Dmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werman T;
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 15-SEP-2000; 2000US-0663870.
 PR 06-NOV-2000; 2000US-0707351.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Dmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werman T;
 XX
 PR WPI: 2001-465571/50.
 DR N-P5DB; AAH94530.
 XX
 PT Novel fetal proteins useful for the treatment and diagnosis of diseases
 PT associated with dysfunction of the protein e.g. cancers, immune
 PT disorders, growth disorders, thrombolytic disorders, nervous system
 PT disorders and inflammation -
 PT disorders, growth disorders, thrombolytic disorders, nervous system
 PT disorders and inflammation -
 PS Example 4; Page 600-601; 715pp; English.
 XX
 CC The invention relates to novel foetal polypeptides encoded by
 CC poly nucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal poly nucleotides and polypeptides are
 CC useful in the treatment and diagnosis of diseases such as cancers,
 CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
 CC disorders, nervous system disorders and inflammation. The present
 CC sequence is a polypeptide encoded by a cDNA assembled using
 CC an expressed sequence tag (EST) found to be expressed in human
 CC foetal tissue cDNA libraries.
 XX
 SQ Sequence 102 AA;
 Query Match 71.4%; Score 45; DB 22; Length 102;
 Best Local Similarity 75.0%; Pred. No. 9.1; Mismatches 6; Conservative 0; Indels 2; Gaps 0;
 Matches 6; Mismatches 0; Gaps 0;
 QY 1 SPRKWWPTC 8
 ID 16 SRRWWPVC 23
 XX
 RESULT 8
 AAM06855 ID AAM06855 standard; Protein; 107 AA.
 XX
 AC AAM06855;
 XX
 DT 05-OCT-2001 (first entry)
 DE Human foetal protein, SEQ ID NO: 1063.
 XX
 KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 KW nootropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antisense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation.
 OS Homo sapiens.
 XX
 PN WO200155339-A2.
 XX
 PD 02-AUG-2001.
 XX
 PR 25-JAN-2001; 2001WO-US02723.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 15-SEP-2000; 2000US-0663870.
 PR 06-NOV-2000; 2000US-0707351.
 XX
 PA (HYSE-) HYSEQ INC.

specification. The foetal polyribonucleotides and polypeptides are useful in the treatment and diagnosis of diseases such as cancers, immune disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders, nervous system disorders and inflammation. The present sequence is a polypeptide encoded by a cDNA assembled from an expressed sequence tag (EST) found to be expressed in human foetal tissue cDNA libraries.

CC	Best Local Similarity	71.4%	Score	45;	DB	22;	Length	107;	
CC	Matches	6;	Pred.	No.	9.5;		Mismatches	2;	Indels
CC	Qy	1 SPRWWPTC	B	Db	16 SRRWWPVC	23	O;	Gaps	0;
CC	SO	Sequence	107 AA;						
XX	RESULT	9							
AAG22709	XX	DT	17-OCT-2000	(first entry)					
ID AAG22709	XX	DB	zea mays	protein fragment SEQ ID NO:	25743.				
AC AAG22709;	XX	KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence; corn.						
XX	OS	OS	zea mays	subsp.	mays.				
PN	EP1033405-A2.	XX							
XX	PD	06-SEP-2000.							
XX	PP	25-FEB-2000;	2000EP-0301439.						
XX	PR	25-FEB-1999;	99US-0121825.						
PR	05-MAR-1999;	99US-0123180.							
PR	09-MAR-1999;	99US-0123348.							
PR	23-MAR-1999;	99US-0125788.							
PR	25-MAR-1999;	99US-0126664.							
PR	29-MAR-1999;	99US-0126785.							
PR	01-APR-1999;	99US-0127462.							
PR	08-APR-1999;	99US-012814.							
PR	16-APR-1999;	99US-0129845.							
PR	19-APR-1999;	99US-0130077.							
PR	21-APR-1999;	99US-0130449.							
PR	23-APR-1999;	99US-0130510.							
PR	06-APR-1999;	99US-0130891.							
PR	28-APR-1999;	99US-0131449.							
PR	30-APR-1999;	99US-0132048.							
PR	04-MAY-1999;	99US-0132407.							
PR	05-MAY-1999;	99US-0132484.							
PR	06-MAY-1999;	99US-0132486.							
PR	06-MAY-1999;	99US-0132863.							
PR	07-MAY-1999;	99US-0134256.							
PR	11-MAY-1999;	99US-0134218.							
PR	14-MAY-1999;	99US-0134219.							
PR	14-MAY-1999;	99US-0134221.							
PR	14-MAY-1999;	99US-0134287.							
PR	18-MAY-1999;	99US-0134370.							
PR	19-MAY-1999;	99US-0134768.							
PR	20-MAY-1999;	99US-0135124.							
PR	21-MAY-1999;	99US-0135353.							
PR	24-MAY-1999;	99US-0135629.							
PR	02-AUG-1999;	99US-0146389.							
PR	25-MAY-1999;	99US-0136021.							
PR	27-MAY-1999;	99US-0136392.							
PR	28-MAY-1999;	99US-0136782.							
PR	01-JUN-1999;	99US-0137222.							
PR	01-JUN-1999;	99US-0137528.							
PR	04-JUN-1999;	99US-0137502.							
PR	07-JUN-1999;	99US-0137724.							
PR	10-JUN-1999;	99US-0138094.							
PR	14-JUN-1999;	99US-0138847.							
PR	16-JUN-1999;	99US-0139119.							
PR	17-JUN-1999;	99US-0139452.							
PR	18-JUN-1999;	99US-0139455.							
PR	18-JUN-1999;	99US-0139456.							
PR	18-JUN-1999;	99US-0139457.							
PR	18-JUN-1999;	99US-0139458.							
PR	18-JUN-1999;	99US-0139460.							
PR	18-JUN-1999;	99US-0139461.							
PR	18-JUN-1999;	99US-0139463.							
PR	18-JUN-1999;	99US-0139750.							
PR	18-JUN-1999;	99US-0139817.							
PR	22-JUN-1999;	99US-0139899.							
PR	23-JUN-1999;	99US-0140354.							
PR	24-JUN-1999;	99US-0140695.							
PR	28-JUN-1999;	99US-0140823.							
PR	29-JUN-1999;	99US-0140991.							
PR	30-JUN-1999;	99US-0141287.							
PR	01-JUL-1999;	99US-0141842.							
PR	01-JUL-1999;	99US-0142154.							
PR	02-JUL-1999;	99US-0142055.							
PR	06-JUL-1999;	99US-0142390.							
PR	08-JUL-1999;	99US-0142803.							
PR	09-JUL-1999;	99US-0142920.							
PR	12-JUL-1999;	99US-0142977.							
PR	13-JUL-1999;	99US-0143542.							
PR	14-JUL-1999;	99US-0143624.							
PR	15-JUL-1999;	99US-0144005.							
PR	16-JUL-1999;	99US-0144085.							
PR	16-JUL-1999;	99US-0144086.							
PR	16-JUL-1999;	99US-0144225.							
PR	19-JUL-1999;	99US-0144331.							
PR	19-JUL-1999;	99US-0144332.							
PR	19-JUL-1999;	99US-0144333.							
PR	19-JUL-1999;	99US-014334.							
PR	19-JUL-1999;	99US-0144355.							
PR	20-JUL-1999;	99US-0144532.							
PR	20-JUL-1999;	99US-0144804.							
PR	21-JUL-1999;	99US-0145086.							
PR	21-JUL-1999;	99US-0145088.							
PR	22-JUL-1999;	99US-0145085.							
PR	22-JUL-1999;	99US-0145087.							
PR	22-JUL-1999;	99US-0145089.							
PR	22-JUL-1999;	99US-0145145.							
PR	23-JUL-1999;	99US-0145088.							
PR	23-JUL-1999;	99US-0145085.							
PR	26-JUL-1999;	99US-0145087.							
PR	27-JUL-1999;	99US-0145084.							
PR	27-JUL-1999;	99US-0145145.							
PR	27-JUL-1999;	99US-0145148.							
PR	28-JUL-1999;	99US-0145086.							
PR	28-JUL-1999;	99US-0145085.							
PR	29-JUL-1999;	99US-0145224.							
PR	29-JUL-1999;	99US-0145226.							
PR	29-JUL-1999;	99US-0145227.							
PR	29-JUL-1999;	99US-0145228.							
PR	29-JUL-1999;	99US-0145229.							
PR	29-JUL-1999;	99US-0145230.							
PR	29-JUL-1999;	99US-0145231.							
PR	29-JUL-1999;	99US-0145232.							
PR	29-JUL-1999;	99US-0145233.							
PR	29-JUL-1999;	99US-0145234.							
PR	29-JUL-1999;	99US-0145235.							
PR	29-JUL-1999;	99US-0145236.							
PR	29-JUL-1999;	99US-0145237.							
PR	29-JUL-1999;	99US-0145238.							
PR	29-JUL-1999;	99US-0145239.							
PR	29-JUL-1999;	99US-0145240.							
PR	29-JUL-1999;	99US-0145241.							
PR	29-JUL-1999;	99US-0145242.							
PR	29-JUL-1999;	99US-0145243.							
PR	29-JUL-1999;	99US-0145244.							
PR	29-JUL-1999;	99US-0145245.							
PR	29-JUL-1999;	99US-0145246.							
PR	29-JUL-1999;	99US-0145247.							
PR	29-JUL-1999;	99US-0145248.							
PR	29-JUL-1999;	99US-0145249.							
PR	29-JUL-1999;	99US-0145250.							
PR	29-JUL-1999;	99US-0145251.							
PR	29-JUL-1999;	99US-0145252.							
PR	29-JUL-1999;	99US-0145253.							
PR	29-JUL-1999;	99US-0145254.							
PR	29-JUL-1999;	99US-0145255.							
PR	29-JUL-1999;	99US-0145256.							
PR	29-JUL-1999;	99US-0145257.							
PR	29-JUL-1999;	99US-0145258.							
PR	29-JUL-1999;	99US-0145259.							
PR	29-JUL-1999;	99US-0145260.							
PR	29-JUL-1999;	99US-0145261.							
PR	29-JUL-1999;	99US-0145262.							
PR	29-JUL-1999;	99US-0145263.							
PR	29-JUL-1999;	99US-0145264.							
PR	29-JUL-1999;	99US-0145265.							
PR	29-JUL-1999;	99US-0145266.							
PR	29-JUL-1999;	99US-0145267.							
PR	29-JUL-1999;	99US-0145268.							
PR	29-JUL-1999;	99US-0145269.							
PR	29-JUL-1999;	99US-0145270.							
PR	29-JUL-1999;	99US-0145271.							
PR	29-JUL-1999;	99US-0145272.							
PR	29-JUL-1999;	99US-0145273.							
PR	29-JUL-1999;	99US-0145274.							
PR	29-JUL-1999;	99US-0145275.							
PR	29-JUL-1999;	99US-0145276.							
PR	29-JUL-1999;	99US-0145277.							
PR	29-JUL-1999;	99US-0145278.							
PR	29-JUL-1999;	99US-0145279.							
PR	29-JUL-1999;	99US-0145280.							
PR	29-JUL-1999;	99US-0145281.							
PR	29-JUL-1999;	99US-0145282.			</				

PR	03-AUG-1999;	99US-014703B.
PR	04-AUG-1999;	99US-0147204.
PR	05-AUG-1999;	99US-0147303.
PR	05-AUG-1999;	99US-0147192.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	16-AUG-1999;	99US-0149369.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154773.
PR	21-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0158393.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160441.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160769.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0161406.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.8%; Score 44; DB 21; Length 136;
Best Local Similarity 75.0%; Pred. No. 9.2.;
Matches 6; Conservative 0; Mismatches 1; Indels 0;
Gaps 0;
QY 2 PRWWPTC 8
Db 24 PRWWTC 30

RESULT 10
AAV25308

ID AAV25308 standard; Protein; 49 AA.

XX

AC AAV25308;

XX

DT 03-SEP-1999 (first entry)

XX

DE HCV NS5B carboxy-terminus protein fragment from genotype HCV J491.

XX

KW NS5B; antiviral compound; truncated protein; mutant; immunoprotective;

vaccine; immunological response; protection; disease; Flaviviridae;

KW virus; x-ray crystallography; hydrophobic tail.

XX Hepatitis C virus.

OS

XX PN WO9929843-A1.

XX PD

XX 17-JUN-1999.

XX 09-DEC-1998; 98WO-US26070.

XX PR 11-DEC-1997; 97US-0069208.

XX PA (SMIK) SMITHKLINE BECHAM CORP.

XX PI Del Vecchio A;

XX DR WPI; 1999-404939/34.

XX PT Nucleic acid encoding Hepatitis C Virus NS5B truncated protein

XX Disclosure; Page 16; 63pp; English.

CC This invention describes a novel nucleic acid encoding a Hepatitis C Virus (HCV) truncation mutant of NS5B. The product of the invention has immunoprotective activity and can be used as a vaccine. The NS5B mutant or a nucleic acid vector directing expression of the mutant, are used to induce an immunological response in a mammal. Antibodies against the NS5B mutant are used to protect mammals against diseases caused by viruses of the Flaviviridae. Additionally, soluble protein produced by the method would allow for determination of the structure of the protein via x-ray crystallography or other known methods. The HCV NS5B truncation mutant has a deletion of a hydrophobic tail which release the protein into the soluble portion of the cell, allowing for a greater recovery of soluble protein for screening for inhibitors of NS5B enzymatic activity. AAV25308-Y25322 represent the carboxy-terminus of the NS5B protein from various HCV genotypes which are used to describe the method of the invention.

CC Sequence 49 AA;

CC Query Match 69.3%; Score 43; DB 20; Length 49;
Best Local Similarity 75.0%; Pred. No. 9.2.;
Matches 6; Conservative 1; Mismatches 1; Indels 0;
Gaps 0;

QY 2 PRWWPTC 9
Db 27 PRWWPLCL 34

RESULT 11

AAB95763 standard; Protein; 237 AA.

XX

AC

XX

ID

XX

AC

XX

DT

XX

DE

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8: SEQ ID 18691; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification, where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAK13653 to AAK13629 represent human cDNA sequences; RAB92446 to AAB95893 represent human amino acid sequences; and AAK13629 to AAK13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 237 AA;

Query Match Similarity 68.3%; Score 43; DB 22; Length 237;
Best Local Similarity 100.0%; Pred No. 40; Mismatches 5; Conservative 0; Indels 0; Gaps 0;

Qy 4 WMPRC 8

Db 99 WMPRC 103

RESULT 12

AAU69468 standard; Protein; 238 AA.

XX

ID AAU69468;

XX

AC AAU69468;

XX

DT 29-JAN-2002 (first entry)

XX

DE Human purified secretory polypeptide #37.

XX

DE Human: purified secretory polypeptide; cell proliferative disorder; cancer; immune system disorder; demyelinating disorder; motor neuron disorder; neurofibromatosis; neuromuscular disorder; central nervous system disorder; enzyme linked immunosorbent assay; ELISA; gene therapy.

XX

OS Homo sapiens.

XX

PA WO200162918-A2.

XX

PD 30-AUG-2001.

XX

PR 01-FEB-2001; 2001WO-US03465.

XX

PR 24-FEB-2000; 2000US-185215P.

PR 24-FEB-2000; 2000US-185216P.

PR 16-MAY-2000; 2000US-205212P.

PR 17-MAY-2000; 2000US-205216P.

PR 17-MAY-2000; 2000US-205217P.

PR 17-MAY-2000; 2000US-205333P.

PR 17-MAY-2000; 2000US-205324P.

XX

PA (INCYT-) INCYTE GENOMICS INC.

XX

PI Panzar SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;

PI Chen A, D'sa SA, Amshay S, Dahl CR, Dam RC, Daniels SE;

PI Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL;

PI Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;

PI Wright RJ, Yao PE, Yu JY, Bradley DL, Bratcher SR, Chen W;

PI Cohen HJ, Hodgson DM, Lincoln SE;

DR WPI; 2001-648217/74.

XX

PT Nucleic acids encoding secretory polypeptides, useful in genetic diagnosis and therapy -

XX

PS Disclosure; Page 211-212; 237pp; English.

XX

Sequences AAU69432-AAU69511 represent purified secretory polypeptides of the invention. The polypeptides and their associated polynucleotides can be used in the treatment, prevention and diagnosis of diseases associated with inappropriate secretory protein expression. These diseases include cell proliferative disorders such as atherosclerosis and psoriasis, cancers such as leukaemia and melanoma, immune system disorders such as asthma and diabetes mellitus, neurological disorders such as epilepsy and Parkinson's disease, mental disorders such as schizophrenia and seasonal affective disorder (SAD), motor neuron disorders such as amyotrophic lateral sclerosis, demyelinating disorders such as multiple sclerosis, central nervous system disorders such as mental retardation and neurofibromatosis and neuromuscular disorders such as cerebral palsy and muscular dystrophy. Target polynucleotides in a sample can be detected by hybridising the sample with a probe sequence complementary to the target polynucleotide, under conditions in which a hybridisation complex is formed, and detecting the presence or absence of the complex. The polypeptides may also be used as antigens in the production of antibodies against secretory proteins and in assays to identify modulators of protein expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the sequences of the invention in samples e.g. by enzyme linked immunosorbent assay (ELISA).

SQ Sequence 238 AA;

DR
DR
N-PSDB; AAV20420.

XX
PT Human discs large 1 gene family - useful in, e.g. therapeutic
composition(s) for treating cancer

XX
PS Claim 8; Page 18-20; 31pp; Japanese.

XX
CC The present sequence represents human dlx-2 molecule. The present
CC invention describes human discs large 1 gene (dlx) family expressible in
CC nerve tissue. Also described are: (1) a polynucleotide (PN) encoding dlx
CC and comprising a 3100 (I) or 3213 bp (II) sequence; (2) RNA corresponding
CC to PN having sequences (I) and (II); (3) RNA having a base sequence
CC translated by dlx; (4) an antisense PN having a at least 15 bp sequence,
CC and which is a part of PN of (1); (5) a derivative of the antisense PN;
(6) an antibody specific for dlx, and (7) an antibody specific for a
CC polypeptide having a 817 or 869 aa sequence. The polypeptide, RNAs and
CC antibodies can be used for detection of dlx. The antisense PN can be
CC used as a therapeutic composition for treating cancer.

XX
SQ Sequence 849 AA;

Query Match 68.3%; Score 43; DB 19; Length 849;

Best Local Similarity 71.4%; Pred. No. 1.3e+02; Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	2	PRWWPPC	8
Db	103	PSWWPEC	109

Search completed: March 27, 2003, 16:10:33
Job time : 6.42105 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

GeneCore version 5.1.4_p5_4578

M protein - protein search, using sw model

run on: March 27, 2003, 16:08:52 ; Search time 1.94737 seconds

(Without alignments)
135,982 Million cell updates/sec

Title: US-10-019-219-2
Effect score: 63
Sequence: 1 SPRWHPTCI 9
Scoring table: BLOSUM62
Gopop: 10.0 , Gapext 0.5
Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Cost-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/POTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match length	DB ID	Description
1	43	68.3	49	4 US-09-208-140-20
2	43	68.3	3010	4 US-09-014-416-3
3	40	63.5	520	1 US-08-261-822A-10
4	40	63.5	520	5 PCT-US95-07744A-10
5	39	61.9	376	2 US-09-758-621-10
6	39	61.9	376	4 US-09-107-858-10
7	39	61.9	567	1 US-08-261-822A-12
8	39	61.9	567	5 PCT-US95-07744A-12
9	38	60.3	584	1 US-08-261-822A-8
10	38	60.3	584	5 PCT-US95-07744A-8
11	38	61.9	1182	4 US-09-287-354-6
12	38	60.3	1189	4 US-09-287-354-3
13	38	60.3	1189	4 US-09-287-354-4
14	38	60.3	1207	4 US-09-287-354-5
15	37	58.7	49	3 US-08-851-843A-216
16	37	58.7	49	4 US-08-974-549A-335
17	37	58.7	49	4 US-08-854-050-216
18	37	58.7	49	4 US-09-430-323-216
19	37	58.7	94	3 US-08-851-843A-215
20	37	58.7	94	4 US-08-974-549A-334
21	37	58.7	94	4 US-08-854-050-215
22	37	58.7	94	4 US-08-974-549A-335
23	37	58.7	311	4 US-09-430-323-216
24	37	58.7	311	4 US-09-277-716-18
25	37	58.7	311	4 US-09-609-161B-18
26	37	58.7	311	4 US-08-808-589A-2
37	58.7	339	4 US-09-330-317B-2	

ALIGNMENTS

RESULT	1	US-09-208-140-20
SEQ_ID	20	Application US/09208140
PATENT_NO	6228576	
GENERAL INFORMATION:		
APPLICANT:	Del Vecchio, Alfred	
TIME OF INVENTION:	HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN	
FILE REFERENCE:	PS0743	
CURRENT APPLICATION NUMBER:	US/09/208,140	
CURRENT FILING DATE:	1998-12-09	
NUMBER OF SEQ ID NOS:	34	
SOFTWARE:	FASTSEQ for Windows Version 3.0	
SEQ_ID NO	20	
LENGTH:	49	
TYPE:	PRT	
ORGANISM:	Viral	
US-09-208-140-20		
Query	Match	
Best	Local Similarity	68.3%;
Matches	Conservative	75.0%;
		Pred. No. 2.3;
		DB 4;
		Length 49;
		1; Mismatches
		1; Indels
		0; Gaps
OY	2	PRWWHPTCI 9
Db	27	PRWFPCII 34
RESULT 2		
US-09-014-416-3		
SEQUENCE	3	Application US/09014416
PATENT NO.	6153421	
GENERAL INFORMATION:		
APPLICANT:	Yanagi, Masayuki	
APPLICANT:	Burk, Jens	
APPLICANT:	Emerson, Susanne U.	
APPLICANT:	Purcell, Robert H.	
TIME OF INVENTION:	CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND	
FILE REFERENCE:	20264276	
CURRENT APPLICATION NUMBER:	US/09/014,416	
CURRENT FILING DATE:	1998-01-27	
EARLIER APPLICATION NUMBER:	US 60/053,062	
EARLIER FILING DATE:	1997-07-18	
NUMBER OF SEQ ID NOS:	65	
SOFTWARE:	PatentIn Ver. 2.1	
SEQ_ID NO	3	
LENGTH:	3010	
TYPE:	PRT	
ORGANISM:	Hepatitis C virus	

US-09-014-416-3

Query Match 68.3%; Score 43; DB 4; Length 3010;

Best Local Similarity 75.0%; Pred. No. 1.3e+02; 1; Mismatches 0; Gaps 0;

Matches 6; Conservative 1; Indels 0; Gaps 0;

RESULT 3
US-08-261-822A-10
Sequence 10, Application US/08261822A
Patent No. 565053

GENERAL INFORMATION:

APPLICANT: Ecker, Joseph R. et al.

TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene

TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene

NUMBER OF SEQUENCES: 82

CORRESPONDENCE ADDRESS:

ADDRESSE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 565053ris

STREET: One Liberty Place, 46th floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07744A

FILING DATE: 15-JUNE-1995

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/261-822

FILING DATE: June 17, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Beardell, Lori Y.

REGISTRATION NUMBER: 34,293

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 520 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLogy: linear

MOLECULE TYPE: Peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

PCT-US95/07744A-10

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07744A

FILING DATE: 15-JUNE-1995

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/261-822

FILING DATE: June 17, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Beardell, Lori Y.

REGISTRATION NUMBER: 34,293

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 520 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLogy: linear

MOLECULE TYPE: Peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

PCT-US95/07744A-10

Query Match 63.5%; Score 40; DB 5; Length 520;

Best Local Similarity 71.4%; Pred. No. 66; Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 5
US-08-758-621-10
Sequence 10, Application US/08758621

GENERAL INFORMATION:

APPLICANT: Guarinot, Mary Lou, and Elde, David J.

TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/758,621

FILING DATE: 29-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Silver, Jean M.

REGISTRATION NUMBER: 39,030

REFERENCE/DOCKET NUMBER: DOI-099CP

TELECOMMUNICATION INFORMATION:

Query Match 63.5%; Score 40; DB 4; Length 3010;

Best Local Similarity 75.0%; Pred. No. 1.3e+02; 1; Mismatches 0; Gaps 0;

Matches 6; Conservative 1; Indels 0; Gaps 0;

RESULT 4
PCT-US95/07744A-10
Sequence 10, Application PC/TU9507744A

GENERAL INFORMATION:

APPLICANT: Trustees of The University of Pennsylvania

TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene

TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene

NUMBER OF SEQUENCES: 82

CORRESPONDENCE ADDRESS:

ADDRESSE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris

STREET: One Liberty Place, 46th floor

ATTONEY/AGENT INFORMATION:
 NAME: Beardell, Lori Y.
 REGISTRATION NUMBER: 34,293
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 376 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

Query Match 61.9%; Score 39; DB 2; Length 376;
 Best Local Similarity 66.7%; Pred. No. 69; Mismatches 3; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRKWWPTCL 9
 Db 277 SKRWWWPWAL 285

RESULT 6
 US-09-107-858-10
 ; Sequence 10, Application US/09107858
 ; Patent No. 6162900
 ; GENERAL INFORMATION:
 APPLICANT: Guerinot, Mary Lou et al.
 TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR
 FILE REFERENCE: DCI-0990CPDV

CURRENT APPLICATION NUMBER: US/09/107,858
 EARLIER APPLICATION NUMBER: 08/758,621
 EARLIER FILING DATE: 1996-11-27
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 10
 LENGTH: 376
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 US-09-107-858-10

Query Match 61.9%; Score 39; DB 4; Length 376;
 Best Local Similarity 66.7%; Pred. No. 69; Mismatches 3; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRKWWPTCL 9
 Db 277 SKRWWWPWAL 285

RESULT 7
 US-08-261-822A-12
 ; Sequence 12, Application US/08261822A
 ; Patent No. 5650553
 ; GENERAL INFORMATION:
 APPLICANT: Ecker, Joseph R. et al.
 TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
 NUMBER OF SEQUENCES: 82
 CORRESPONDENCE ADDRESS:
 ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT-US95/07744A
 FILING DATE: 15-JUNE-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/261,822
 FILING DATE: June 17, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Beardell, Lori Y.
 REGISTRATION NUMBER: 34,293
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 567 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 PCT-US95/07744A-12

Query Match 61.9%; Score 39; DB 5; Length 567;
 Best Local Similarity 83.3%; Pred. No. 1e+02; Mismatches 5; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRKWWPTCL 9
 Db 277 SKRWWWPWAL 285

RESULT 9
 US-08-261-822A-8
 Sequence 8, Application US/08261822A
 Patent No. 5650553
 GENERAL INFORMATION:
 APPLICANT: ECKER, Joseph R. et al.
 TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
 TITLE OF INVENTION: and Pathogens
 NUMBER OF SEQUENCES: 82
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & NO. 5650553ris
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, version #1.25
 CURRENT APPLICATION DATA:
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/261,822A
 FILING DATE: 17-JUN-1994
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Beardell, Lori Y.
 REGISTRATION NUMBER: 34,293
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEX/FAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 584 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ;
 US-08-261-822A-8

Query Match Score 38; DB 1; Length 584;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservatve 0; Mismatches 1; Indels 0; Gaps 0;

QY	1 SPRWWP 6
Db	213 SPPWWP 218

RESULT 10
 PCT-US95-07744A-8
 Sequence 8, Application PC/TUS9507744A
 GENERAL INFORMATION:
 APPLICANT: Trustees of The University of Pennsylvania
 TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
 TITLE OF INVENTION: and Pathogens
 NUMBER OF SEQUENCES: 82
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:

Query Match Score 38; DB 4; Length 1182;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 6; Conservatve 1; Mismatches 2; Indels 0; Gaps 0;

QY	1 SPRWWPCL 9
Db	155 APFWWPCL 163

RESULT 11
 US-09-287-354-6
 Sequence 6, Application US/09287354
 Patent No. 6348348
 GENERAL INFORMATION:
 APPLICANT: THOMPSON, Catherine C.
 TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
 FILE REFERENCE: Thompson-20263/024/435
 CURRENT APPLICATION NUMBER: US/09/287,354
 CURRENT FILING DATE: 1999-04-07
 CURRENT APPLICATION NUMBER: US 60/080,888
 EARLIER FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 5
 LENGTH: 1182
 TYPE: PRT
 ORGANISM: Mouse

Query Match Score 38; DB 5; Length 584;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservatve 0; Mismatches 1; Indels 0; Gaps 0;

QY	1 SPRWWP 6
Db	213 SPPWWP 218

RESULT 12
 US-09-287-354-6
 Sequence 3, Application US/09287354
 Patent No. 6348348
 GENERAL INFORMATION:
 APPLICANT: THOMPSON, Catherine C.
 TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN

FILE REFERENCE: Thompson-2063/0243435
 CURRENT APPLICATION NUMBER: US/09/287,354
 CURRENT FILING DATE: 1999-04-07
 EARLIER APPLICATION NUMBER: US 60/080,888
 EARLIER FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 3
 LENGTH: 1189
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-287-354-3

Query Match 60.3%; Score 38; DB 4; Length 1189;
 Best Local Similarity 66.7%; Pred. No. 3e+02; 1; Mismatches 2; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWWPTCL 9
 Db 155 APFWWPCTCL 163

RESULT 13
 US-09-287-354-4
 Sequence 4, Application US/09287354
 Patent No. 634348
 GENERAL INFORMATION:
 APPLICANT: THOMPSON, Catherine C.
 TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
 FILE REFERENCE: Thompson-2023/0243435
 CURRENT APPLICATION NUMBER: US/09/287,354
 CURRENT FILING DATE: 1999-04-07
 EARLIER APPLICATION NUMBER: US 60/080,888
 EARLIER FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 4
 LENGTH: 1189
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-287-354-4

Query Match 60.3%; Score 38; DB 4; Length 1189;
 Best Local Similarity 66.7%; Pred. No. 3e+02; 1; Mismatches 2; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWWPTCL 9
 Db 155 APFWWPCTCL 163

RESULT 14
 US-09-287-354-5
 Sequence 5, Application US/09287354
 Patent No. 6348348
 GENERAL INFORMATION:
 APPLICANT: THOMPSON, Catherine C.
 TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
 FILE REFERENCE: Thompson-20263/0243435
 CURRENT APPLICATION NUMBER: US/09/287,354
 CURRENT FILING DATE: 1999-04-07
 EARLIER APPLICATION NUMBER: US 60/080,888
 EARLIER FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 5
 LENGTH: 1207
 TYPE: PRT
 ORGANISM: Rat
 US-09-287-354-5

Query Match 60.3%; Score 38; DB 4; Length 1207;
 Best Local Similarity 66.7%; Pred. No. 3e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SPRWWPTCL 9
 Db 181 APFWWPCTCL 189

RESULT 15
 US-08-851-843A-216
 Sequence 216, Application US/08851843A
 Patent No. 6093009
 GENERAL INFORMATION:
 APPLICANT: Cech, Thomas R.
 APPLICANT: Ligner, Joachim
 APPLICANT: Nakamura, Toru
 APPLICANT: Chapman, Karen B.
 APPLICANT: Morin, Greg B.
 APPLICANT: Harley, Calvin
 APPLICANT: Andrews, William H.
 TITLE OF INVENTION: NO. 6093809el. Telomerase
 NUMBER OF SEQUENCES: 225
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94111
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/851,843A
 FILING DATE: 06-MAY-1997
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/844,419
 FILING DATE: 18-APR-1997
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/724,643
 FILING DATE: 01-OCT-1996
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-002930US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 216:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 49 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-851-843A-216

Query Match 58.7%; Score 37; DB 3; Length 49;
 Best Local Similarity 80.0%; Pred. No. 19;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WWPTC 8
 Db 3 WWPEC 7

Mon Apr 7 09:26:00 2003

Search completed: March 27, 2003, 16:13:23
Job time : 2.94737 secs

us-10-019-219-2.rai

GenCore version 5.1.4-p5-4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

Om protein - protein search, using sw model
 Run on: March 27, 2003, 16:12:13 ; Search time 2.15789 Seconds
 (without alignments) 244.921 Million cell updates/sec

Title: US-10-019-219-2
 Perfect score: 63
 Sequence: 1 SPRWWPTCL 9
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 237916 seqs, 58723674 residues
 Total number of hits satisfying chosen parameters: 237916
 Minimum DB seq length: 0
 Maximum DB seq length: 200000000
 Post-processing: Minimum Match 0%, Maximum Match 100%
 Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pupbaa/US08_NEW_PUB_pep: *
 2: /cgn2_6/ptodata/2/pupbaa/PTC_NEW_PUB_pep: *
 3: /cgn2_6/ptodata/2/pupbaa/US06_NEW_PUB_pep: *
 4: /cgn2_6/ptodata/2/pupbaa/US06_PUBCOMB_pep: *
 5: /cgn2_6/ptodata/2/pupbaa/US07_NEW_PUB_pep: *
 6: /cgn2_6/ptodata/2/pupbaa/US07_PUBCOMB_pep: *
 7: /cgn2_6/ptodata/2/pupbaa/PC08_PUBCOMB_pep: *
 8: /cgn2_6/ptodata/2/pupbaa/US08_PUBCOMB_pep: *
 9: /cgn2_6/ptodata/2/pupbaa/US09_NEW_PUB_pep: *
 10: /cgn2_6/ptodata/2/pupbaa/US09_PUBCOMB_pep: *
 11: /cgn2_6/ptodata/2/pupbaa/US10_NEW_PUB_pep: *
 12: /cgn2_6/ptodata/2/pupbaa/US10_PUBCOMB_pep: *
 13: /cgn2_6/ptodata/2/pupbaa/US60_NEW_PUB_pep: *
 14: /cgn2_6/ptodata/2/pupbaa/US60_PUBCOMB_pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	46	73.0	72	10 US-09-764-869-648
2	43	68.3	49	10 US-09-733-183A
3	42	66.7	34	10 US-09-864-761-42036
4	41	65.1	68	9 US-09-980-480-449
5	40	63.5	155	10 US-09-729-674-104
6	39	61.9	77	10 US-09-864-761-3634
7	38	60.3	1189	9 US-10-024-368-6
8	38	60.3	1189	9 US-10-024-368-4
9	38	60.3	1207	9 US-10-024-368-5
10	38	60.3	1207	9 US-10-020-704-319
11	37.5	59.5	135	9 US-10-050-704-325
12	37.5	59.5	177	9 US-10-050-704-316
13	37.5	59.5	177	9 US-10-050-704-11732
14	37	59.5	279	10 US-09-815-247-11732
15	37	58.7	17	10 US-09-864-761-46894
16	37	58.7	49	9 US-09-843-676-216
17	37	58.7	49	9 US-09-430-486-216
18	37	58.7	49	9 US-10-053-758-216
19	37	58.7	49	9 US-10-054-295-216

ALIGNMENTS

RESULT 1
 US-09-764-869-648
 ; Sequence 648, Application US/09764869
 ; Patent No. US20061521A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC007
 ; CURRENT APPLICATION NUMBER: US/09/764,869
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 24420
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 648
 ; LENGTH: 72
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-09-764-869-648

Query Match 73.0%; Score 46; DB 10; Length 72;
 Best Local Similarity 85.7%; Pred. No. 2.5; Mismatches 0; Indels 0; Gaps 0;

QY	3 RWWPTCL 9	Db	49 RWWPCL 55
RESULT 2 US-09-733-183A-20			
; Sequence 20, Appli			
; Sequence 42036, A			
; Sequence 449, Appli			
; Sequence 104, Appli			
; Sequence 36434, A			
RESULT 2 US-09-733-183A-20			
; Sequence 20, Appli			
; Patent No. US20061521A1			
; GENERAL INFORMATION:			
; APPLICANT: Delvecchio, Alfred M.			
; TITLE OF INVENTION: Hepatitis C Virus NS5B Truncated Protein			
; FILE REFERENCE: P50743D1			
; CURRENT APPLICATION NUMBER: US/09/733,183A			
; CURRENT FILING DATE: 2000-12-08			
; PRIOR APPLICATION NUMBER: 09/208140			
; PRIOR FILING DATE: 1998-12-09			
; PRIOR APPLICATION NUMBER: 60/069208			

PRIOR FILING DATE: 1997-12-11
 NUMBER OF SEQ ID NOS: 34
 SOFTWARE: FastSEQ for Windows version 3.0
 SEQ ID NO 20
 LENGTH: 49
 TYPE: PRI
 ; ORGANISM: Viral
 ; US-09-733-183n-20

Query Match 68.3%; Score 43; DB 10; Length 49;
 Best Local Similarity 75.0%; Pred. No. 4.9;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWWPTCL 9
 Db 27 PRWFPLCL 34

RESULT 3
 US-09-864-761-42036

; Sequence 42036, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:

; APPLICANT: Penn, Sharona G.
 ; APPLICANT: Rank, David R.

; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aeomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-05-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Amaxax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 42036
 ; LENGTH: 34

TYPE: PRN

ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC017004.2

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.3
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 12
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.9
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.8
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 14

Query Match 66.7%; Score 42; DB 10; Length 34;
 Best Local Similarity 83.3%; Pred. No. 4.9;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWWPCTC 8
 Db 19 RWPPC 24

RESULT 4
 US-09-986-480-449

; Sequence 449, Application US/09986480
 ; Publication No. US20030027999A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 143 Human Secreted Proteins

; FILE REFERENCE: P5500P1
 ; CURRENT APPLICATION NUMBER: US/09/985,480
 ; CURRENT FILING DATE: 2001-11-08
 ; PRIOR APPLICATION NUMBER: PCT/US00/12788
 ; PRIOR FILING DATE: 2000-05-11
 ; PRIOR APPLICATION NUMBER: US 60/134,068
 ; PRIOR FILING DATE: 1999-05-13
 ; NUMBER OF SEQ ID NOS: 456
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 449
 ; LENGTH: 68

; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (28)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-09-986-480-449

Query Match 65.1%; Score 41; DB 9; Length 68;
 Best Local Similarity 100.0%; Pred. No. 13;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWWP 6
 Db 6 PRWWP 10

RESULT 5
 US-09-729-674-104

; Sequence 104, Application US/09729674
 ; Patent No. US20010039335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: LaVaille, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treaty, Maurice
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Steininger II, Robert J.
 ; APPLICANT: Spaulding, Vicki

APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Rechtein, Kim
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 FILE REFERENCE: 6055-64X
 CURRENT APPLICATION NUMBER: US/09/729,674
 PRIOR APPLICATION NUMBER: 09/539,330
 PRIOR FILING DATE: 2000-12-04
 NUMBER OF SEQ ID NOS: 283
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 104
 LENGTH: 155
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: UNSURE
 LOCATION: (46)
 ; US-09-729-674-104
 Query Match Score 40; DB 10; Length 155;
 Best Local Similarity 63.5%; Pred. No. 36; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0;
 Oy 2 PRMWPTCL 9
 | | |||
 Db 61 PAWQACL 68
 RESULT 6
 US-09-864-761-36434
 Sequence 36434, Application US/09864761
 ; Patent No. US2002048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aecomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/1180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; SEQ ID NO 6
 ; LENGTH: 1182
 ; TYPE: PRT
 ; ORGANISM: Mouse
 ; US-10-024-368-6
 Query Match Score 38; DB 9; Length 1182;
 Best Local Similarity 60.3%; Pred. No. 4.2e+02; Mismatches 2; Indels 0; Gaps 0;
 Matches 6; Conservative 1;
 Oy 1 SPRMWPTCL 9
 ; US-10-024-368-6
 Db 155 :|||111| 163
 RESULT 8
 US-10-024-368-3
 ; Sequence 3, Application US/10024368
 ; Publication No. US20030027300A1

GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: THOMPSON-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIORITY NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-024-368-3

Query Match ; Best Local Similarity 60.3%; Score 38; DB 9; Length 1189;
; Sequence 4, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: THOMPSON-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-04-07
; PRIORITY NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-024-368-4

RESULT 9
US-10-024-368-4
; Sequence 4, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: RUBEN ET AL.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 319
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-050-704-319

Query Match ; Best Local Similarity 60.3%; Score 38; DB 9; Length 1189;
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: THOMPSON-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIORITY NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
; US-10-024-368-5

Query Match ; Best Local Similarity 66.7%; Score 38; DB 9; Length 1207;
; Sequence 319, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: RUBEN ET AL.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 319
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-050-704-319

RESULT 11
US-10-050-704-319
; Sequence 319, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: RUBEN ET AL.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 319
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-050-704-319

Query Match ; Best Local Similarity 60.0%; Score 37.5; DB 9; Length 36;
; Sequence 325, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: RUBEN ET AL.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 325
; LENGTH: 325

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-050-704-325

Query Match      59.5%; Score 37.5; DB 9; Length 135;
Best Local Similarity 60.0%; Pred. No. 74; Matches 6; Conservative 0; Mismatches
QY   2 PRW--WPTC 8
    ||| -|| | |
Db   34 PRWTWLWPIC 43

RESULT 13
US-10-050-704-316
Sequence 316, Application US/10050704
Publication No. US20030050442A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.

TITLE OF INVENTION: 62 Human Secreted Proteins
FILE REFERENCE: PZ039P1
CURRENT APPLICATION NUMBER: US/10/050,704
CURRENT FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: 09/684,524
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: PCT/US00/08979
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: 60/128,693
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 60/130,991
PRIOR FILING DATE: 1999-04-26
NUMBER OF SEQ ID NOS: 344
SEQ ID NO 316
LENGTH: 177
TYPE: PRT
ORGANISM: Homo sapiens

US-10-050-704-316

Query Match      59.5%; Score 37.5; DB 9; Length 177;
Best Local Similarity 60.0%; Pred. No. 94; Matches 6; Conservative 0; Mismatches
QY   2 PRW--WPTC 8
    ||| -|| | |
Db   76 PRWTWLWPIC 85

RESULT 14
US-09-815-242-11732
Sequence 11732, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haebelbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELTRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578

RESULT 15
US-09-864-761-46894
Sequence 46894, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-09-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203

RESULT 16
US-09-815-242-11732
Query Match      59.5%; Score 37.5; DB 10; Length 279;
Best Local Similarity 58.3%; Pred. No. 1.4e+02; Matches 7; Conservative 0; Mismatches
QY   1 SPRW--WPTC 9
    ||| -|| | |
Db   17 SPWWWRWCWPCCL 28

; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-815-242-11732

Query Match      59.5%; Score 37.5; DB 10; Length 279;
Best Local Similarity 58.3%; Pred. No. 1.4e+02; Matches 7; Conservative 0; Mismatches
QY   1 SPRW--WPTC 9
    ||| -|| | |
Db   17 SPWWWRWCWPCCL 28

```

PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46894
LENGTH: 17
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO ALL18499.5

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.95

Query Match Similarity 58.7%; Score 37; DB 10; Length 17;
Best Local Similarity 80.0%; Pred. No. 14;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 WWPTC 8
Db 4 WWPTC 8

Search completed: March 27, 2003, 16:20:26
Job time : 3.15789 secs

Run on:	March 27, 2003, 16:01:27 ;	Search time 1.52632 Seconds (without alignments)	
copyright (c) 1993 - 2003 Compugen Ltd.		566.862 Million cell updates/sec	
OM protein - protein search, using sw model			
Title:	US-10-019-219-2		
Perfect score:	63		
Sequence:	1 SPRWWPPCL 9		
Scoring table:	BLOSUM62		
Searched:	283224 seqs, 96134422 residues		
Total number of hits satisfying chosen parameters:	283224		
Minimum DB seq length: 0			
Maximum DB seq length: 200000000			
Post-processing: Minimum Match 0%			
Database : PIR 73;*		Maximum Match 100%	
1: pir1;*		Listing first 45 summaries	
2: pir2;*			
3: pir3;*			
4: pir4;*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
		SUMMARIES	
Result No.	Score	Query Match Length DB ID	Description
1	46	73.0	806 2 G84863
2	43	68.3	120 2 T42054
3	42	66.7	221 2 C81099
4	42	66.7	221 2 C81842
5	42	66.7	275 2 T40211
6	42	66.7	443 2 139538
7	42	66.7	2092 2 S30026
8	42	66.7	2149 2 S18676
9	41	65.1	123 2 H72698
10	41	65.1	252 1 J00417
11	41	65.1	295 2 AD3577
12	41	65.1	461 2 H70899
13	41	65.1	790 2 T34293
14	41	65.1	1398 2 T18350
15	40	63.5	64 2 S28486
16	40	63.5	212 2 S74376
17	40	63.5	263 1 S43189
18	40	63.5	290 1 D47468
19	40	63.5	471 2 T50016
20	40	63.5	542 2 I39540
21	40	63.5	997 2 T39521
22	39	61.9	119 2 A53257
23	39	61.9	176 2 H75332
24	39	61.9	338 2 A87566
25	39	61.9	376 2 S33654
26	39	61.9	567 2 E96764
27	38.5	61.1	423 2 AG2394
28	38	60.3	2 S26718
29	38	60.3	175 2 T27543
			ALIGNMENTS
			RESULT 1
			G84863
			hypothetical protein At2g43240 [imported] - Arabidopsis thaliana
			C;Species: Arabidopsis thaliana (mouse-ear cress)
			C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
			C;Accession: G84863
			R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.W.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Sher, M.; Vanken, S.E.; Umayam, L.; Tallon, J.; Niemzner, W.C.; White, O.; Eisen, J.J.A.; Salzberg, S.L.; Fraser, C.M.; Venetian Nature 402, 761-768, 1999
			A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
			A;Reference number: A84420; MUID:20083487; PMID:1017197
			A;Accession: G84863
			A;Status: preliminary
			A;Molecule type: DNA
			A;Residues: 1-806 <STO>
			A;Cross-references: GB:AB002093; NID:9763933; PIDN: AAC64313.1; GSPDB:GN00139
			A;Genes: At2g43240
			A;Map position: 2
			RESULT 2
			Query Match Best Local Similarity 73.0%; Score 46; DB 2; Length 806; Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
	QY	2 PRWWPPCL 9	
	Db	676 PKWWPPSL 683	
			T4054
			hypothetical protein - Streptomyces coelicolor (fragment)
			C;Species: Streptomyces coelicolor
			C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
			C;Accession: T42054
			R;Guillarro, J.; Santamaría, R.; Schauer, A.; Losick, R.
			J; Bacteriol. 170, 1895-1901, 1988
			A;nitE: Promoter determining the timing and spatial localization of transcription of the nitE gene
			A;Reference number: Z22043; MUID:8816921; PMID:450872
			A;accession: T42054
			A;Status: preliminary; translated from GB/EMBL/DDBJ
			A;Molecule type: DNA
			A;Residues: 1-120 <GUI>
			A;Cross-references: EMBL:M20145; PIDN:AAA26812.1
			Query Match Best Local Similarity 68.3%; Score 43; DB 2; Length 120; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
	QY	4 WWPPCL 8	

Db 88 RWWPTC 92

RESULT 3

C81099 hypothetical protein **NMB1296** [imported] - *Neisseria meningitidis* (strain MC58 serogroup C;Species: *Neisseria meningitidis*
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: C81099
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haff, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; ri, H.; Qin, H.; Vaithianathan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; ve
A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: C81099
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-221 <BET>
A;Cross-references: GB:AE002478; GB:AE002098; NID:97226533; PIDN:AAF41672.1; PID:9722653
A;Experimental source: serogroup B, strain MC58
A;Genetics:
A;Gene: NMB1296

Query Match Score 66.7%; Score 42; DB 2; Length 221;
Best Local Similarity 83.3%; Pred. No. 21; 1; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 RWWPRC 8
Db 177 RWWPOC 182

RESULT 4

C81842 hypothetical protein **NMA1506** [imported] - *Neisseria meningitidis* (strain 22491 serogroup C;Species: *Neisseria meningitidis*
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: C81842
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel ; Holdroyd, S.; Jagels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-505, 2000
A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491.
A;Reference number: A811775; MUID:20222556; PMID:10761919
A;Accession: C81842
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-221 <PAR>
A;Cross-references: GB:AL162756; GB:AL157959; NID:97380091; PIDN:CABB4738.1; PID:9738015
A;Experimental source: serogroup A, strain 22491
A;Genetics:
A;Gene: NMA1506

RESULT 5

A;Accession: IX0211
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: DNA
A;Residues: 1-275 <RES>
A;Cross-references: EMBL:U12678; NID:g529961; PIDN: AAC28892.1; PID:g529965
C;Superfamily: oxidoreductase
C;Keywords: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology <SDADH>
F;7-190/Domain: short-chain alcohol dehydrogenase homology <SDADH>
Query Match Score 66.7%; Score 42; DB 2; Length 275;
Best Local Similarity 71.4%; Pred. No. 25; 1; Mismatches 5; Conservative 1; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 PRWWPTC 8
Db 219 PRWWRSC 225

RESULT 6

139538 alpha amylase - *Aeromonas hydrophila*
C;Species: *Aeromonas hydrophila*
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 03-Dec-1999
C;Accession: 139538
R;Chang, M.C.; Chang, J.C.; Chen, J.P. J. Gen. Microbiol. 139, 3215-223, 1993
A;Title: Cloning and nucleotide sequence of an extracellular alpha-amylase gene from
A;Reference number: 139538; MUID:94172314; PMID:8126440
A;Accession: 139538
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: DNA
A;Residues: 1-443 <RES>
A;Cross-references: GB:L19299; NID:9304014; PIDN:AAA21016.1; PID:9304015
A;Genetics:
C;Genetics:

F;166-291/Domain: alpha-amylase core homology <AMY>
Query Match Score 66.7%; Score 42; DB 2; Length 443;
Best Local Similarity 83.3%; Pred. No. 39; 1; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 7

S30026 genome polyprotein - Rift Valley fever virus
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: Rift Valley fever virus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S30026
R;Mueller, R.; Argentini, C.; Bouloy, M.; Prehaud, C.; Bishop, D.H.L.
A;Title: coriendum; Completion of the genome sequence of Rift Valley fever phlebovi
A;Reference number: S30026
A;Accession: S30026
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-2092 <NUL>
A;Cross-references: GB:X56464
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992
C;Keywords: nucleotidyltransferase

RESULT 8

Query Match Score 66.7%; Score 42; DB 2; Length 2092;
Best Local Similarity 83.3%; Pred. No. 1.5e-02; 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SPRWWP 6
Db 177 RWWPOC 182

RESULT 9

A;Accession: S30026

C;Species: *Bradyrhizobium japonicum*

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 08-Oct-1999

C;Accession: I40211

R;Tuili, R.E.; Kistler, D.L.

Appl. Environ. Microbiol. 59, 4136-4142, 1993

A;Title: Cloning and mutagenesis of a cytochrome P-450 locus from *Bradyrhizobium japonicum*

A;Reference number: I40207

Query Match 65.1%; Score 41; DB 2; Length 1398;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWWWP 6
 |||||
 Db 373 PRWWWP 377

RESULT 13

T34293 hypothetical protein F49E10.2 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Accession: T34293
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 R;Miller, N.
 submitted to the EMBL Data Library, April 1996
 A;Description: The sequence of *C. elegans* cosmid F49E10.
 A;Reference number: Z21500
 A;Accession: T34293
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-190 <ML>
 A;Cross-references: EMBL:U53341; PIDN: AAC69106.1; GSPPDB:GN00028; CESP:F49E10.2
 A;Experimental source: strain Bristol N2; clone F49E10
 C;Genetics:
 A;Gene: CESP:F49E10.2
 A;Map position: X
 A;Introns: 51/1; 92/3; 176/3; 235/3; 332/2; 514/1; 543/2; 569/3; 677/1; 732/3

Query Match 65.1%; Score 41; DB 2; Length 790;
 Best Local Similarity 100.0%; Pred. No. 91; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWWWP 6
 |||||
 Db 688 PRWWWP 692

RESULT 14

T18350 probable pol polyprotein - rice blast fungus gypsy retroelement (fragment)
 C;Species: *Magnaporthe grisea* (rice blast fungus)
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T18350
 R;Dobinson, K.F.
 submitted to the EMBL Data Library, September 1994
 A;Description: Sequence of the gry retroelement.
 A;Reference number: Z18803
 A;Accession: T18350
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1398 <DOB>
 A;Cross-references: EMBL:M77661; NTD:g538065; PID:g538067; PIDN:AAA21442.1
 C;Genetics:
 A;Mobile element: gypsy retroelement

Query Match 65.1%; Score 41; DB 2; Length 1398;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWWWP 6

Db 76 PRWWWP 80

RESULT 15

S28486 hypothetical protein 2 - *Vibrio cholerae*
 C;Species: *Vibrio cholerae*
 C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
 C;Accession: S28486
 R;Manning, P.A.
 submitted to the EMBL Data Library, May 1991
 A;Reference number: S28467
 A;Accession: S28467
 A;Molecule type: DNA
 A;Residues: 1-64 <MAN>
 A;Cross-references: EMBL:X59554; NID:g48381; PIDN:CAA42152.1; PID:g48401
 A;Experimental source: strain 017

Query Match 63.5%; Score 40; DB 2; Length 64;
 Best Local Similarity 75.0%; Pred. No. 14; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRWWPTC 8
 |||||
 Db 51 SYRWPTC 58

Search completed: March 27, 2003, 16:12:41
 Job time: 3.52632 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.4-p5-4578

P24115 salmonella

P36435 sacccharomy

P37385 synechococc

P23989 streptococc

P19711 bovine vira

P87379 xenopus lae

P25640 sacccharomy

Q918P6 mycoplasma

P84890 haemophilus

P58184 caulosbacter

P24327 caenorhabdi

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:00:22 ; Search time 1.52632 Seconds

(without alignments) (244.567 Million cell updates/sec)

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMWPCTCL 9

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 10%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	43	68.3	817	1 DLG3_HUMAN
2	43	68.3	849	1 DLG3_RAT
3	42	66.7	275	1 YCPI_BRUA
4	42	66.7	443	1 AMVA_AERB
5	42	66.7	849	1 DLG3_MOUSE
6	42	66.7	249	1 RRPL_RVRYZ
7	41	65.1	252	1 SAST_VIRAN
8	41	65.1	461	1 C12_MICRU
9	39	61.9	119	1 ATP6_NASFO
10	39	61.9	176	1 ZRTL_YEAST
11	38	60.3	175	1 YOCA_CAEEL
12	38	60.3	211	1 LYCH_CHASP
13	38	60.3	374	1 KILL_SCHEO
14	38	60.3	379	1 METY_PSEAE
15	38	60.3	417	1 OSPE_DROME
16	38	60.3	1181	1 HAIR_RAT
17	38	60.3	1182	1 HAIR_MOUSE
18	38	60.3	1189	1 HAIR_HUMAN
19	37	58.7	309	1 ER25_YEAST
20	37	58.7	311	1 LUCI_RENFE
21	37	58.7	418	1 VIAN_SHEEP
22	37	58.7	460	1 YS15_MYCTU
23	37	58.7	546	1 LNT_TRPA
24	37	58.7	809	1 YATA_SCHPO
25	37	58.7	1131	1 YANG_SCHPO
26	37	58.7	1385	1 YM55_CAEPD
27	36.5	57.9	830	1 HMTL_SCHPO
28	36	57.1	94	1 YOR6_NMV
29	36	57.1	283	1 ISPE_CHIMD
30	36	57.1	333	1 A85G_MYCTU
31	36	57.1	340	1 A85C_MYCTU
32	36	57.1	352	1 A85C_MYCAV
33	36	57.1	431	1 CINI_ECOLI

ALIGNMENTS

RESULT 1	DLG3_HUMAN	STANDARD;	PRY;	817 AA.
ID	DLG3_HUMAN			
AC	Q92796	homo sapien		
DT	01-NOV-1997	(Rel. 35, Created)		
DT	01-NOV-1997	(Rel. 35, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	(Neuroendocrine-DIG)	(NE-DIG) (Discs, large homolog 3).		
GN	DLG3 OR KIAA1232.			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OC	OX NCBI_TAXID=606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=9732623; PubMed=9188857;			
RA	Makino K., Kuwahara H., Masuko N., Nishiyama Y., Morisaki T., Sasaki J., Nakao M., Kuwano A., Nakata M., Ushio Y., Sayo H.;			
RA	P41131 aeromonas h			
RA	P70175 mus musculu			
RA	P27316 rift valley			
RA	P19829 vibrio angu			
RA	P77900 mycobacteri			
RA	P22067 naegleria f			
RA	P32804 saccharomy			
RA	Q23280 caenorhabdi			
RA	P00721 chalaropsis			
RA	O14300 schizosacch			
RA	P57714 pseudomonas			
RA	P01924 drosophila			
RA	P97609 rattus norv			
RA	Q61645 mus musculu			
RA	Q43593 homo sapien			
CC	P33045 saccharomy			
CC	P27652 renilla ren			
CC	P48043 ovis aries			
CC	P010809 mycobacteri			
CC	Q83279 treponema p			
CC	Q01055 schizosacch			
CC	Q01077 schizosacch			
CC	P4501 caenorhabdi			
CC	Q02592 schizosacch			
DR	EMLB; U49089; AAB61453.1; -			
DR	EMBL; AB033058; BA86546.1; -			
DR	HSSP; Q12989; 1PDR.			
DR	Genew; HGNC;902; DLG3.			
DR	MMR; 300189; -			
DR	InterPro; IPR000619; Guanylate_kin.			
DR	InterPro; IPR01478; PDZ.			

DR	InterPro; IPR001452; SH3.	CC	modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR	Pfam; PF00018; SH3; 1.	CC	entities or send an email to license@isb-sib.ch .
DR	Pfam; PF00595; PDZ; 3.	CC	-----
DR	ProDom; PD00625; Guanylate_kin; 1.	DR	-----
DR	SMART; SM00072; GuKc; 1.	DR	-----
DR	SMART; SM00228; PDZ; 3.	DR	HSSP; Q12959; IPR00619; Guanylate_kin.
DR	PROSITE; PS00556; SH3; 1.	DR	InterPro; IPR001478; PDZ.
DR	PROSITE; PS50052; GUANYLATE_KINASE_2; 1.	DR	InterPro; IPR001452; SH3.
DR	PROSITE; PS50002; SH3; 1.	DR	SMART; SM00072; GuKc; 1.
DR	PROSITE; PS50028; PDZ; 3.	DR	SMART; SM00326; SH3; 1.
KW	SH3 domain; Repeat;	DR	PROSITE; PS50052; Guanylate_kin; 1.
FT	DOMAIN 130; PDZ 1.	DR	PRODOM; PD00006; SH3; 1.
FT	DOMAIN 226; PDZ 2.	DR	SMART; SM00228; PDZ; 3.
FT	DOMAIN 379; PDZ 3.	DR	SMART; SM00072; GuKc; 1.
FT	DOMAIN 503; SH3.	DR	PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
FT	DOMAIN 628; 803; CONFLICT 330; 381; FTALADNHITHNNSGGVAVGKSYSPAPQPVPTRYSI PRMLAEEDG-> AARRRGAMERKARYPSGLAMGLGS ASASAWRRAQRWRWPLRSRPGGA (IN REF. 2).	DR	PROSITE; PS50052; Guanylate_kin; 1.
FT	CONFFLICT 592; 606; ENMAGESSSTEQGVNL-> SIKTKRKSFPLSKRPPFYKSK	DR	PROSITE; PS50052; PDZ; 3.
FT	SEQUENCE 817 AA; 90344 MW; 3D7512E4713FC4E CRC64;	FT	SEQUENCE 817 AA; 90344 MW; 3D7512E4713FC4E CRC64;
Qy	Query Match 68.3%; Score 43; DB 1; Length 817; Best Local Similarity 71.4%; Pred. No. 22; Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	FT	SEQUENCE 817 AA; 90344 MW; 3D7512E4713FC4E CRC64;
Db	2 PRWWPTC 8 Db 103 PSWWPEC 109	FT	SEQUENCE 817 AA; 90344 MW; 3D7512E4713FC4E CRC64;
RESULT 2			
ID	DLG3_RAT	ID	DLG3_RAT
ID	STANDARD;	PRT;	849 AA.
AC	062236; P70547;	AC	062236; P70547;
DT	01-Nov-1997 (Rel. 35, Created)	DT	01-Nov-1997 (Rel. 35, last sequence update)
DT	15-JUN-2002 (Rel. 41, last annotation update)	DE	presynaptic protein SAP102 (synapse-associated protein 102) (PSD-95/SAP90 related protein 1) (Discs, large homolog 3).
GN	DLG3.	DB	GN
OS	Rattus norvegicus (Rat).	OC	Rattus norvegicus
OC	Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TAXID=10116;	RN	[1] SEQUENCE FROM N.A. (LONG FORM).
RN	SEQUENCE FROM N.A. (LONG FORM).	RC	TISSUE=Brain;
RX	MEDLINE-B6374358; PubMed=8786649;	RA	MEDLINE-B6374358; PubMed=8786649;
RA	Mueller B.M., Kistner U., Kindlner S., Chung W.J., Kuhlendahl S., Fenster S.D., Lau L.-F., Veh R.W., Huguenir R.L., Gundelfinger E.D., Garner C.C.; "SAP102, a novel postsynaptic protein that interacts with NMDA receptor complexes in vivo.", Neuron 17:255-265(1996).	RA	"SAP102, a novel postsynaptic protein that interacts with NMDA receptor complexes in vivo.", Neuron 17:255-265(1996).
RL	[2]	RA	[2]
RP	SEQUENCE FROM N.A. (SHORT FORM).	RA	SEQUENCE FROM N.A. (SHORT FORM).
RA	Irie M., Haga Y., Takagi Y.; Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.	RA	Tully R.E., Keister D.L.; "Cloning and mutagenesis of a cytochrome P-450 locus from Bradyrhizobium japonicum that is expressed anaerobically and symbiotically."
RL	-1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR	RT	Bradyrhizobium japonicum that is expressed anaerobically and symbiotically."
CC	SUBUNIT NB2B.	RT	AppL Environ.; Microbiol. 59:4136-4142(1993).
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.	RL	[2]
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.	RN	SEQUENCE FROM N.A.
CC	-1- SIMILARITY: CONTAINS 1 GUANYLATE KINASE-LIKE DOMAIN.	RP	SEQUENCE FROM N.A.
CC	-1- SIMILARITY: BELONGS TO THE MACUK FAMILY OF CELL JUNCTION PROTEINS.	RC	STRAIN=USA 110;
CC	-1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR	RX	MEDLINE-B8322110; PubMed=9655913;
CC	SUBUNIT NB2B.	RX	Tully R.E., van Berkum P., Lovins K.W., Keister D.L.; "Identification and sequencing of a cytochrome P450 gene cluster from Bradyrhizobium japonicum."
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.	RT	Bradyrhizobium japonicum.
CC	-1- SIMILARITY: CONTAINS 1 GUANYLATE KINASE-LIKE DOMAIN.	RL	Biophys. Acta 1398:243-255(1998).
CC	-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.	CC	-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U12678; AAC28892.1; -.
 DR HSSP; P08074; ICD.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PRO0080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Hypothetical protein; Oxidoreductase.
 FT NP_BIND; 10 34 NAD OR NADP (BY SIMILARITY).
 ACT_SITE 159 159 BY SIMILARITY.
 SEQUENCE 275 AA; 28827 MW; 312E7070C404DF86 CRC64;
 Query Match 66.7%; Score 42; DB 1; Length 443;
 Best Local Similarity 83.3%; Pred. No. 18; Mismatches 0;
 Matches 5; Conservative 0; MisMatches 1; Indels 0; Gaps 0;
 QY 2 PRKWPCTC 8
 Db 219 PRKWRSC 225

RESULT 4
 ANA_AERRY STANDARD; PRT; 443 AA.
 ID ANA_AERRY STANDARD; PRT; 443 AA.
 AC P41131; 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 RP 01-FEB-1995 (Rel. 31, Last annotation update)
 RC Alpha_amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolyase).
 DE ANA_AERRY.
 OS Aeromonas hydrophila.
 OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
 OC Aeromonas.
 OX NCBI_TaxID=644;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MCC-1;
 RA Chang M.C., Chang J.-C., Chen J.-P.;
 RT Cloning and nucleotide sequence of an extracellular alpha-amylase
 gene from Aeromonas hydrophila MCC-1.";
 RL J. Gen. Microbiol. 139:3215-3223(1993)
 CC -I CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U19299; AAA21016.1; -.
 DR HSSP; P29957; IADM.
 DR InterPro; IPR00061; Alpha_amylase.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Pfam; PF02806; alpha-amylase_C; 1.
 DR PRINTS; PRO0110; ALPHAMYLASE.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 443 ALPHA-AMYLASE.
 FT ACT_SITE 198 198 BY SIMILARITY.
 FT ACT_SITE 202 202 BY SIMILARITY.
 FT ACT_SITE 287 287 BY SIMILARITY.

RESULT 5
 DLG3_MOUSE STANDARD; PRT; 849 AA.
 ID DLG3_MOUSE STANDARD; PRT; 849 AA.
 AC P70175; 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pre-synaptic protein SAP102 (Synapse-associated protein 102) (Discs,
 DE large homolog 3).
 DE DLG3 OR DLGH3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Eutheria; Rodentia; Scurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Kohmura N., Makino S., Yagi T.;
 RL Submitted (Aut-1996) to the EMBL/GenBank/DBJ databases.
 CC -I FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR
 CC SUBUNIT NR2B.
 CC -I SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
 CC -I SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -I SIMILARITY: CONTAINS 1 GUANYLATE KINASE-LIKE DOMAIN.
 CC -I SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D87117; BAA13249.1; -.
 DR HSSP; Q12959; IPR001478; PDZ.
 DR InterPro; IPR001455; SH3.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00595; PDZ; 3.
 DR Pfam; PF00625; Guanylate_kin; 1.
 DR PRODOM; PD000066; SH3; 1.
 DR SMART; SM0072; GUKC; 1.
 DR SMART; SM00220; PDZ; 3.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00856; GUANYLATE_KINASE_1; 1.
 DR PROSITE; PS00052; GUANYLATE_KINASE_2; 1.
 DR PROSITE; PS50010; PDZ; 3.
 DR PROSITE; PS50006; SH3; 1.
 KW SH3 domain; Repeat.
 FT DOMAIN 149 235 PDZ 1.
 FT DOMAIN 244 330 PDZ 2.
 FT DOMAIN 404 484 PDZ 3.
 FT DOMAIN 519 589 SH3.
 FT DOMAIN 659 849 GUANYLATE KINASE.
 SQ SEQUENCE 849 AA; 934862 MW; EF3EF2D7513338EE CRC64;

Query Match 66.7%; Score 42; DB 1; Length 849;
 Best Local Similarity 71.4%; Pred. No. 31; Mismatches 2; Indels 0; Gaps 0;

QY	2	PWMPTC 8
RC		STRAIN-531A;
RX		MEDLINE-93328275; PubMed=8335354;
RA		Tolmasy M.E., Actis L.A., Crosa J.H.;
RT		"A single amino acid change in Angr, a protein encoded by Pjm1-like virulence plasmids, results in hyperproduction of anguibactin.";
RL		Infect. Immun. 61:3228-3233(1993).
CC		-1- FUNCTION: PROBABLE THIOESTERASE INVOLVED IN THE BIOSYNTHESIS OF ANGUBACTIN, AN IRON-BINDING SIDEKOPHORE.
CC		-1- SIMILARITY: TO OTHER THIOESTERASES.
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC		EMBL: M34504; AAAT9861.1; -
DR		EMBL; Z12000; CAR79845.1; -.
DR		PIR; J00417; J00417.
DR		PIR; S26422; S26422.
DR		InterPro; IPR001031; Thioesterase.
DR		Pfam; PF00975; Thioesterase_1.
KW		Plasmid; Hydrolase.
FT		ACT-SITE 92 92 BY SIMILARITY.
FT		ACT-SITE 229 229 BY SIMILARITY.
SQ		SEQUENCE 252 AA; 28070 MW; 1FB1A3CCEDB9F4 CRC64;
QY	1	SPRWMPCL 9
Db	156	SPRWMPFL 164
RESULT 7		
SAST_VIBAN	STANDARD;	PRT; 252 AA.
ID	C132_MCTU	STANDARD;
AC	P19829;	PRT; 461 AA.
DT	01-FEB-1991 (Rel. 17, Created)	
DT	01-FEB-1991 (Rel. 17, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
OS	Vibrio anguillarum (Listonella anguillarum).	
OC	Plasmid pJML.	
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Listonella.	
OC	NCBI_TaxID=55601;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-H37RV;	
RX	MEDLINE-9825987; PubMed=9634230;	
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglemeier K., Gas S., Barry C.E. III, Tekla F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moul S., Murphy L., Oliver S., Osborne J., Quail M.A., Rejkardameen M.A., Rogers J., Rutter S., Seeger K., Skeleton S., Squares S., Squares R., RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;	
RT	"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." Nature 393:537-544(1998).	
RT	[2]	
RN	SEQUENCE FROM N.A.	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-CDC 1551 / Oshkosh;	
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Gene 86:45-51(1990).	
RA	[2]	
RC	SEQUENCE FROM N.A.	
RX	MEDLINE=90185247; PubMed=2311935;	
RA	Farrell D.H., Mikesell P., Actis L.A., Crosa J.H.;	
RT	"A regulatory gene, angr, of the iron uptake system of <i>Vibrio anguillarum</i> ; similarity with phage P22 cro and regulation by iron." Gene 86:45-51(1990).	
RN	SEQUENCE FROM N.A.	

RA Bishai W.;
 RT "whole genome comparison of *Mycobacterium tuberculosis* clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

CC EMBL; 280108; CAB02176.1; -.
 DR EMBL; AB007015; ARK45704.1; -.
 DR HSSP; P14779; LJPZ.
 DR TIGR; MT1439; -.
 DR InterPro; IPR011128; Cytochrome_P450.
 DR Pfam; PF00057; p450; 1.
 DR PRINTS; PR03885; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Hypothetical protein; Oxidoreductase; Monoxygenase; Heme;
 KW Complete proteome;
 FT BINDING 409 HEME (BY SIMILARITY).
 FT CONFLICT 135 R -> L (IN REF. 2).
 SQ SEQUENCE 461 AA; 52229 MW; 2DEF61C8A10B0CF3 CRC64;
 Query Match 65.1%; Score 41; DB 1; Length 119;
 Best Local Similarity 100.0%; Pred. No. 26; Indels 461;
 Matches 5; Conservative 0; Mismatches 1; Gaps 0;
 QY 2 PRWP 6
 |||||
 Db 373 PRWP 377

RESULT 9
 ATP6_NAEFO STANDARD; PRT; 119 AA.
 ID ATP6_NAEFO
 AC P22057;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6) (Fragment).
 GN ATP6 OR OL12.
 OS Naegleria fowleri.
 OG Mitochondrion.
 OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
 OX NCBI_TaxID=5763;
 RN [1]
 SEQUENCE FROM N.A.
 RP SPRAIN-LEE;
 RX MEDLINE=91178040; PubMed=2007628;
 RA McLaughlin G.L., Wodkin M.H., Huizinga H.W.;
 RT "Amplification of repetitive DNA for the specific detection of
 Neagleria fowleri.;"
 RL Clin. Microbiol. 29: 227-230(1991).
 CC -!- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
 DIRECT ROLE IN THE TRANSPORTATION OF PROTONS ACROSS THE MEMBRANE.
 CC -!- FUNCTION: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 H(+)-OUT);
 CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) HAS FIVE
 CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -!- SUBCELLULAR LOCATION: Integral membrane Protein.
 CC -!- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

CC EMBL; M55009; CAB25936.1; -.
 DR PIR; A53257; A53257.
 DR InterPro; IPRO00588; ATPsynt_Asub.
 DR Pfam; PF00119; ATP-synt_A; 1.
 DR PROSITE; PS00449; ATPASE_A; PARTIAL.
 KW Hydrogen ion transport; CF(0); Mitochondrion; transmembrane.
 FT NON_TER 119 119
 SEQUENCE 119 AA; 13334 MW; 50892FC0BB5C04F5 CRC64;
 SQ Query Match 61.9%; Score 39; DB 1; Length 119;
 Best Local Similarity 83.3%; Pred. No. 15; Indels 0; Mismatches 1; Gaps 0;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 WWPPICL 9
 |||||
 Db 70 WWATCL 75

RESULT 10
 ZRT1_YEAST STANDARD; PRT; 376 AA.
 ID ZRT1_YEAST
 AC P3204;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-Nov-1997 (Rel. 35, Last annotation update)
 DE ZRT1 protein
 GN ZRT1 OR YGL255W OR NRC376.
 OC Saccharomyces cerevisiae (Baker's yeast).
 OC Sacccharomyces cerevisiae; Sacccharomycotina; Sacccharomycetes;
 OC Sacccharomycetales; Sacccharomyctaceae; Sacccharomycetes;
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WZ03;
 RT MEDLINE=9311123; PubMed=8322518;
 RA Breitwieser W., Price C., Schuster T.;
 RT "Identification of a gene encoding a novel zinc finger protein in
 Saccharomyces cerevisiae.";
 RL Yeast 9:551-556(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=7127827; PubMed=8912578;
 RA Colissac E., Maillier E., Robineau S., Netter P.;
 RT Sequence of a 39,411 bp DNA fragment covering the left end of
 chromosome VII of *Saccharomyces cerevisiae*.;
 RL Yeast 12:1555-1562(1996).
 CC -!- FUNCTION: HIGH AFFINITY ZINC TRANSPORT PROTEIN.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: TO YEAST YLR130C AND S.POMBE SPc16p10.06.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

CC EMBL; X67787; CAAM7997.1; -.
 DR EMBL; X64357; CAA6132.1; -.
 DR PIR; S2777; CAA96975.1; -.
 DR SGD; S0003224; ZRT1.
 DR InterPro; IPR04698; ZIP_transport..
 DR InterPro; IPR003689; Zn_trnpnprt_zip..

DR Pfam; PF02535; Zip; 1.
 DR TIGRFAMS; TIGR00820; zip; 1.
 KW Transmembrane; Transport; Zinc.
 FT TRANSEM_51; 71; POTENTIAL.
 FT TRANSEM_81; 101; POTENTIAL.
 FT TRANSEM_131; 151; POTENTIAL.
 FT TRANSEM_221; 241; POTENTIAL.
 FT TRANSEM_283; 303; POTENTIAL.
 FT TRANSEM_315; 335; POTENTIAL.
 SEQUENCE 376 AA; 41581 MW; TA1F8367D49BAC3C CRC64;

Query Match 61.9%; Score 39; DB 1; Length 376;
 Best Local Similarity 66.8%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 SPRWWPTCL 9
 Db 277 SKRWWPWL 285

RESULT 11
 YOCA_CAEEL STANDARD; PRT; 175 AA.
 AC 023280;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Hypothetical 19.4 kDa protein ZC395.10 in chromosome III.
 GN ZC395.10
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorida; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderaidae; Caenorhabditidae;
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Connell M..
 RL Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
 CC -I- SIMILARITY: BELONGS TO THE P23 / WOS2 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U13642; AAG0038.1; .
 DR WormBase; ZC395.10; CE01436.
 KW Hypothetical protein.
 FT DOMAIN_145; 175; ASP/GLU-RICH.
 FT DOMAIN_165; 168; POLY-GLU.
 SEQUENCE 175 AA; 19431 MW; D5C136F0446E37A CRC64;

Query Match 60.3%; Score 38; DB 1; Length 175;
 Best Local Similarity 55.6%; Pred. No. 30;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPRWWPTCL 9
 Db 81 TPAWWPRLL 89

RESULT 12
 LYCH_CHASP STANDARD; PRT; 211 AA.
 AC P00721;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1994 (Rel. 01, Last sequence update)
 DE N,O-diacylglucosaminidase (EC 3.2.1.-) (Lysozyme CH).
 OS Chalaropsis sp.

Query Match 61.9%; Score 39; DB 1; Length 211;
 Best Local Similarity 60.3%; Pred. No. 38;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SPRWWPTCL 9
 Db 140 NPSWNSSC 147

RESULT 13
 OX11_SCHPO STANDARD; PRT; 374 AA.
 AC 014300;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome oxidase biogenesis protein oxal-1, mitochondrial precursor.
 GN OXA1-1 OR SPAC01.04.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4965;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RX MEDLINE=20177828; PubMed=10712694;
 RA Bonneroy N., Kermorvant M., Groudinsky O., DuJardin G.;
 RT "The respiratory gene Oxa1 has two fission yeast orthologues which
 RT together encode a function essential for cellular viability.";
 RL Mol. Microbiol. 35:1135-1145(2000).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RX MEDLINE=21848401; Pubmed=1859360;
 RA Wood V., Gwilliam R., Ralbandream M.A., Lyne M., Lyne R., Stewart A.,
 SGouros J., Peat N., Haynes S., Baker S., Basham D., Bowman S.,
 RA Brools K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Ridalgo J., Hodgson G.,

RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	FT	TRANSMEM	335	355	6 (POTENTIAL).
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	FT	DOMAIN	356	357	EXTRACELLULAR (POTENTIAL).
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	FT	TRANSMEM	368	388	7 (POTENTIAL).
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	FT	DOMAIN	389	417	CYTOPLASMIC (POTENTIAL).
RA	Brandon R.C., Rogers Y.-H.R., Blazej R.G., Champé M., Pfeiffer B.D.,	FT	CARBOHD	39	39	N-LINKED (GLCNAC . .) (POTENTIAL).
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,	FT	CARBOHD	157	157	N-LINKED (GLCNAC . .) (POTENTIAL).
RA	Abrial J.F., Agbayani I., An H.-J., Andrews-pranckoch C., Baldwin D.,	FT	SEQUENCE	410	417	de Pablo B., Belcher A., Deng Z., May S.A.D., Dew I., Dietz S.M.,
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	FT	SEQUENCE	417 AA:	47212 MW:	Beeson K.Y., Benos P.V., Beaman B.P., Bhandari D., Bolshakov S.,
RA	Borkovka D., Botchan M.R., Bouck J., Broststein P., Brottier P.,	FT	SEQUENCE	417 AA:	47212 MW:	Borkovka D., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Burkhardt K., Cherry J.M., Cowley S., Dahlke C., Davenport L.B., Davies P.,	FT	SEQUENCE	417 AA:	47212 MW:	Cherry J.M., Cowley S., Dahlke C., Davenport L.B., Davies P.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,	FT	SEQUENCE	417 AA:	47212 MW:	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegham C.,	FT	SEQUENCE	417 AA:	47212 MW:	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	FT	SEQUENCE	417 AA:	47212 MW:	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Durbin R.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	FT	SEQUENCE	417 AA:	47212 MW:	Lasko P., Lei Y., Levitt K.A., Li J., Li Z., Liang Y., Lin X.,
RA	Fosler C., Gabrielson A.E., Gang N.S., Gelbart W.M., Glasser K.,	FT	SEQUENCE	417 AA:	47212 MW:	Glodek A., Gong F., Gorreli J.H., Gu Z., Guan P., Harris M.,
RA	Bortnick A., Gotoh T., McIntosh T.C., McLeod P., McPherson D.,	FT	SEQUENCE	417 AA:	47212 MW:	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Merkulov G., Milshina N.V., Kobarry C., Morris J., Mostrefi A.,	FT	SEQUENCE	417 AA:	47212 MW:	Mount S.M., Moy M., Murphy B., Murphy L., Muniz D.M., Nelson D.L.,
RA	Mount S.M., Nelson K.A., Nixon D.R., Pacieb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	FT	SEQUENCE	417 AA:	47212 MW:	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	FT	SEQUENCE	417 AA:	47212 MW:	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirkas R., Tector C., Turner R., Ventter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock Bach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venet J.C.;
RA	"The genome sequence of <i>Drosophila melanogaster</i> ,"; RT	FT	SEQUENCE	417 AA:	47212 MW:	"The genome sequence of <i>Drosophila melanogaster</i> ,";
RL	Science 287:2185-2195(2000).	FT	SEQUENCE	417 AA:	47212 MW:	RT
CC	-I- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT	CC	SEQUENCE	417 AA:	47212 MW:	RECEPTOR.
CC	-I- RECEPTOR.	CC	SEQUENCE	417 AA:	47212 MW:	-I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC	-I- TISSUE SPECIFICITY: EXPRESSED IN 15% OF THE 120 SENSORY NEURONS	CC	SEQUENCE	417 AA:	47212 MW:	-I- TISSUE SPECIFICITY: EXPRESSED IN 15% OF THE 120 SENSORY NEURONS
CC	WITHIN THE MAXILLARY PALP.	CC	SEQUENCE	417 AA:	47212 MW:	WITHIN THE MAXILLARY PALP.
CC	-I- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED	CC	SEQUENCE	417 AA:	47212 MW:	-I- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED
CC	RECEPTORS.	CC	SEQUENCE	417 AA:	47212 MW:	RECEPTORS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	CC	SEQUENCE	417 AA:	47212 MW:	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	CC	SEQUENCE	417 AA:	47212 MW:	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	European Bioinformatics Institute. There are no restrictions on its	CC	SEQUENCE	417 AA:	47212 MW:	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial	CC	SEQUENCE	417 AA:	47212 MW:	entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch).
CC	or send an email to licensee@isb-sib.ch).	CC	SEQUENCE	417 AA:	47212 MW:	or send an email to licensee@isb-sib.ch).
DR	EMBL: AF127922; AAP26357; 1;	DR	SEQUENCE	417 AA:	47212 MW:	DR
DR	EMBL: AE003679; AAF4256; 1;	DR	SEQUENCE	417 AA:	47212 MW:	DR
DR	FLYBase: FBgn0026399; Or85e.	DR	SEQUENCE	417 AA:	47212 MW:	DR
DR	InterPro: IPR004117; 7tm_6.	DR	SEQUENCE	417 AA:	47212 MW:	DR
DR	InterPro: IPR000276; GPCR_Rhodopsn.	DR	SEQUENCE	417 AA:	47212 MW:	DR
DR	Pfam: PF03949; 7tm_6; 1.	DR	SEQUENCE	417 AA:	47212 MW:	DR
KW	Transmembrane G-protein coupled receptor; Olfaction; Glycoprotein;	KW	SEQUENCE	417 AA:	47212 MW:	KW
KW	Multigene family.	KW	SEQUENCE	417 AA:	47212 MW:	KW
DOMAIN	1 60 EXTRACELLULAR (POTENTIAL).	DOMAIN	SEQUENCE	417 AA:	47212 MW:	DOMAIN
TRANSMEM	1 61 81 1 (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	82 98 CYTOPLASMIC (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	99 119 2 (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	120 159 EXTRACELLULAR (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	160 180 3 (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	181 212 4 (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	213 233 4 (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	234 286 EXTRACELLULAR (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	287 307 5 (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM
TRANSMEM	308 334 CYTOPLASMIC (POTENTIAL).	TRANSMEM	SEQUENCE	417 AA:	47212 MW:	TRANSMEM

Search completed: March 27, 2003, 16:11:10
 Job time : 3.52632 secs

Qy	2 PRWWP 6	DB	48 PKWWP 52
Best Local Similarity	60.3%; Score 38;	Matches	4; Conservative
80.0%; Pred. No. 64;	1; Mismatches	0;	Indels 0; Gaps 0;

NYCSTIK -> FIRIDILIFLYNSRRAVHIVTAGKVVMDV
 NLRSLSVITQASFLTLIQKLAKKTTSEL (IN REF. 1).

GenCore version 5.1.4.p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

Om protein - protein search, using sw model

Run on: March 27, 2003, 16:01:02 ; Search time 2.63158 Seconds

(without alignments) updates/sec 704.681 Million cell

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMMPTCL 9

Scoring table: BloSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL_21;*

1: sp_archea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rabbit:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Length DB ID Description

1 46 73.0 806 10 Q9ZW71 Q9zw71 arabidopsis

2 45 71.4 717 10 Q9XR33 Q9xf33 oryza sativ

3 43 68.3 136 10 Q940S8 Q940S8 rosa hybrid

4 43 68.3 139 2 Q9EXK6 Q9EXK6 planobispor

5 43 68.3 145 2 Q929T7 Q929T7 hepatitis c

6 43 68.3 145 12 Q9R895 Q9R895 hepatitis c

7 43 68.3 145 12 Q72123 Q72123 hepatitis c

8 43 68.3 145 12 Q72244 hepatitis c

9 43 68.3 145 12 Q72125 hepatitis c

10 43 68.3 145 12 Q72126 hepatitis c

11 43 68.3 237 4 Q9HFW0 Q9HFW0 homo sapien

12 43 68.3 383 10 Q9FWG2 Q9fwg2 oryza sativ

13 43 68.3 458 10 Q9RJ1 Q9rj1 oryza sativ

14 43 68.3 615 10 Q9A8P3 cucumis mel

15 43 68.3 3010 12 Q92969 hepatitis c

16 43 68.3 3010 12 Q92970 hepatitis c

ALIGNMENTS

RESULT 1	ID Q9ZW71	Description PRELIMINARY;	PRN: 806 AA.
	ID Q9ZW71;	AC Q9ZW71;	
	DT 01-MAY-1999	DT 01-MAY-1999 (TREMBUREL, 10, Last sequence update)	
	DT 01-DEC-2001	DT 01-DEC-2001 (TREMBUREL, 19, Last annotation update)	
	DE AT:Q943240	DE AT:Q943240 protein.	
	GN ATCG43240.	GN Arabidopsis thaliana (mouse-ear cress).	
	OS Eukaryota; Viriplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	OS Arabidopsis thaliana (mouse-ear cress).	
	OC NCBI_TaxID:3702;	OC	
	RN [1]	RN	
	RP SEQUENCE FROM N.A.	RP	
	RC STRAIN=CV; COLUMBIA;	RC	
	RX MEDLINE=20083487; PubMed=10617197;	RX	
	RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D., Fuji C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V., Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S., Crokin L.A., Shen M., Vanaken M., Umayam L., Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M., Venter J.C., RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana," RT RNL Sequence 402:761-768(1999). RN [2] RP SEQUENCE FROM N.A.	RA	
	RA STRAIN=CV; COLUMBIA;	RA	
	RA Lin X.; RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases. DR EMBL; AC004450; AAC64313.1; "-" SQ SEQUENCE 806 AA; 89071 MW; 1F4D5ED6CBFD89A CRC64;	RA	
	Query Match Best Local Similarity 73.0%; Score 46; DB 10; Length 806; Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		

		Db	106 SPPWWPT 112
OY	2 PRWWPFC 9		
	:		
Db	676 PKWWPTSL 683		
RESULT 2		RESULT 4	
09XF33	PRELIMINARY;	09EX86	PRELIMINARY;
ID 09XF33	PRT; 717 AA.	ID 09EX86	PRT; 139 AA.
AC 09XF33;		AC 09EX86;	
DT 01-NOV-1999 (TREMBLrel. 12, Created)		DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)		DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DE Hypothetical 79.0 kDa protein.		DE Putative peptide synthetase (Fragment).	
OS Oryza sativa (indica cultivar-group).		OS Planobispora rosea.	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		OC Bacteria; Firmicutes; Actinobacteria; Actinomycetaceae; Streptosporangiaeae; Planobispora.	
OC Spermato phyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		OC	
OC Ehrhartoideae; Oryzeae; Oryza.		OX NCBI_TaxID=35762;	
OX NCBI_TaxID=39946;		RN [1]	
RN [1]		RP SEQUENCE FROM N.A.	
RP SEQUENCE FROM N.A.		RC STRAIN=ATCC 53733;	
RC STRAIN=CV_ TEQING;		RX MEDLINE=2035709; PubMed=11085259;	
RA Liaca V., Lou A., Young S., Messing J.;		RA SOSIO M., Bossi E., Bianchi A., Donadio S.;	
RT "Microcollinearity in cereal genomes.";		RT "Multiple peptide synthetase gene clusters in actinomycetes.";	
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.		RL Mol. Gen. Genet. 264:213-221(2000).	
DR EMBL; AF128457; AAD27632; 1.		DR EMBL; AJ216363; CAC01622; 1; -.	
DR Interpro; IPR000210; BTB_PoZ.		DR I4SP; P14607; IANU.	
DR Pfam; PF00651; BTB; 1.		DR Interpro; IPR00873; AMP-bind.	
DR SMART; SM00225; BTB; 1.		DR Pfam; PF00501; AMP-binding; 1.	
DR PROSITE; PS50097; BTB; 1.		DR Pfam; PF00550; PP-binding; 1.	
KW Hypothetical protein.		DR PROSITE; PS50075; ACE_DOMAIN; 1.	
KW SEQUENCE 717 AA; 79014 MW; 90DBB88B78119B60 CRC64;		FT NON_TER 1 1 1	
Query Match 71.4%; Score 45; DB 10; Length 717;		FT NON_TER 139 139	
Best Local Similarity 71.4%; Pred. No. 38;		FT NON_TER 139 MW; 1F2489785FD715C6 CRC64;	
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		SQ SEQUENCE 139 AA; 15029 MW;	
QY 2 PRWWPFC 8		Query Match 68.3%; Score 43; DB 2; Length 139;	
Db 302 PKWWPTSL 308		Best Local Similarity 85.7%; Pred. No. 19;	
QY 1 SPRWWPT 7		Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Db 31 SPGWPT 37			
RESULT 3		RESULT 5	
0940SB	PRELIMINARY;	092977	PRELIMINARY;
ID 0940SB	PRT; 136 AA.	ID 092977	PRT; 145 AA.
AC 0940SB;		AC 092977; 092978;	
DT 01-DEC-2001 (TREMBLrel. 19, Created)		DT 01-NOV-1998 (TREMBLrel. 09, Created)	
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)		DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)	
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)		DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE EIN3-like transcription factor (Fragment).		DE Genome polyprotein (Fragment).	
GN EIN3.		OS Hepatitis C virus.	
OS Rosa hybrid cultivar.		OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		OC Hepacivirus.	
OC Spermato phyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		OC NCBITaxID=1103;	
OC eudicots 1; Rosales; Rosaceae; Rosoidea; Rosa.		RN [1]	
OX NCBI_TaxID=128735;		RP SEQUENCE FROM N.A.	
RN [1]		RC STRAIN=HC_J4;	
RP SEQUENCE FROM N.A.		RX MEDLINE=9824094; PubMed=9581788;	
RA Mueller R., Owen C.A., Stummwoll B.M.;		RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,	
RT "Partial sequence of Rosa hybrida cultivar mRNA for EIN3-like transcription factor," to the EMBL/GenBank/DBJ databases.		RA Buchk J.;	
RT Submitted (AUG-2001); to the EMBL/GenBank/DBJ databases.		RT "Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo."	
DR EMBL; AY052825; AAC14267; 1; -.		RL Virology 244:161-172(1998).	
FT NON_TER 1 1		DR EMBL; AF034260; AAC15733; 1; -.	
FT 136 136		DR EMBL; AF054261; AAC15734; 1; -.	
SQ SEQUENCE 136 AA; 15267 MW; B77FAD5PBE1383B9 CRC64;		DR EMBL; AF034265; AAC15738; 1; -.	
Query Match 68.3%; Score 43; DB 10; Length 136;		DR InterPro; IPR002166; HCV_RdRp.	
Best Local Similarity 85.7%; Pred. No. 19;		DR Pfam; PF00999; HCV_RdRp; 1.	
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		FT NON_TER 1 1	
OY 1 SPRWWPT 7		FT NON_TER 145 145	
		FT SEQUENCE 145 AA; 16235 MW; 8157D290205C2252 CRC64;	

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWPTCL 9
 ||::||
 Db 123 PRWFPLCL 130

RESULT 6

ID QW895 PRELIMINARY; PRT; 145 AA.

AC QW895; 01-NOV-1999 (TREMBREL. 12, Created)
 DT 01-NOV-1999 (TREMBREL. 12, Last sequence update)

DE Genome Polyprotein (Fragment).

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OX Hepacivirus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HC-J4;

RX MEDLINE=98240944; PubMed=9581788;

RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 Burk J.;

RT "Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo"; Virology 244:161-172(1998); EMBL: AF054267; AAC15736.1; -.

DR InterPro: IPR002166; HCV_RdRp.

PFam: PF00998; HCV_RdRp; 1.

KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.

FT NON_TER 145 145

SEQUENCE 145 AA; 16175 MW; 8157D280215C2252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWPTCL 9
 ||::||
 Db 123 PRWFPLCL 130

RESULT 7

ID 072123 PRELIMINARY; PRT; 145 AA.

AC 072123; 01-AUG-1998 (TREMBREL. 07, Created)
 DT 01-AUG-1998 (TREMBREL. 07, Last sequence update)

DE Genome polyprotein (Fragment).

OS Hepatitis C virus.

OX Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

NCBI_TaxID=1103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HC-J4;

RX MEDLINE=98240944; PubMed=9581788;

RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 Burk J.;

RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are infectious in vivo"; Virology 244:161-172(1998); EMBL: AF054262; AAC15735.1; -.

DR InterPro: IPR002166; HCV_RdRp; 1.

PFam: PF00998; HCV_RdRp; 1.

RESULT 8

ID 072124 PRELIMINARY; PRT; 145 AA.

AC 072124; 01-AUG-1998 (TREMBREL. 07, Created)
 DT 01-AUG-1998 (TREMBREL. 07, Last sequence update)

DE Genome Polyprotein (Fragment).

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OX Hepacivirus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HC-J4;

RX MEDLINE=98240944; PubMed=9581788;

RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 Burk J.;

RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are infectious in vivo"; Virology 244:161-172(1998); EMBL: AF054267; AAC15736.1; -.

DR InterPro: IPR002166; HCV_RdRp.

PFam: PF00998; HCV_RdRp; 1.

KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.

FT NON_TER 145 145

SEQUENCE 145 AA; 16037 MW; 83AAFD4C3C9DBE252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWPTCL 9
 ||::||
 Db 123 PRWFPLCL 130

RESULT 9

ID 072125 PRELIMINARY; PRT; 145 AA.

AC 072125; 01-AUG-1998 (TREMBREL. 07, Created)
 DT 01-AUG-1998 (TREMBREL. 07, Last sequence update)

DE Genome polyprotein (Fragment).

OS Hepatitis C virus.

OX Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

NCBI_TaxID=1103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HC-J4;

RX MEDLINE=98240944; PubMed=9581788;

RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 Burk J.;

RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are infectious in vivo"; Virology 244:161-172(1998); EMBL: AF054262; AAC15735.1; -.

DR InterPro: IPR002166; HCV_RdRp; 1.

PFam: PF00998; HCV_RdRp; 1.

DR EMBL; AF054266; AAC15739_1; -.
 DR InterPro; IPR002156; HCV_RdRP.
 DR Pfam; PF00998; HCV_RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
 FT NON_TER 145 145
 SEQUENCE 145 AA; 16175 MW; 8157C2CD799E252 CRC64;
 Query Match Similarity 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 PRWMPCTL 9
 Db 123 PRWFPLCL 130

RESULT 10
 072126 PRELIMINARY; PRT; 145 AA.
 ID 072126; DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Genome polyprotein (fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=1103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=HC-J4;
 RX MEDLINE=9224094;
 RA Yanagi M., St.Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 RT infectious in vivo.";
 RL Virology 244:61-172 (1998).
 DR EMBL; AF054268; AAC15741_1;
 DR InterPro; IPR002156; HCV_RdRP.
 DR Pfam; PF00998; HCV_RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
 FT NON_TER 145 145
 SEQUENCE 145 AA; 16175 MW; 8157C2CD799E252 CRC64;
 Query Match Similarity 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 PRWMPCTL 9
 Db 123 PRWFPLCL 130

RESULT 11
 09H7W0 PRELIMINARY; PRT; 237 AA.
 ID 09H7W0; AC 09H7W0;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE CDNA FLJ14202 f1s, clone NT2R3302985.
 OC Homo sapiens (Human).
 OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,

RESULT 12
 09FWG2 PRELIMINARY; PRT; 383 AA.
 ID 09FWG2; AC 09FWG2;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 411.9 kDa protein.
 GN OSJNBB001511_26.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV; NIPPONBARE;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hisiao J., Zisman V., Pal G., Bowman C.L., Fuji C.Y., VanAken S.E.,
 RA Bowman C.L., Craven B., Utterback R.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNBB001511 genomic sequence.";
 RL Submitted (SEP 2001) to the EMBL/GenBank/DDJB databases.
 DR EMBL; AC051633; AAC13599_1;
 DR IISPP; P35813; IAG60.
 DR InterPro; IPR00222; PP2C.
 DR InterPro; IPR01932; PP2C-like.
 DR Pfam; PF00481; PP2C; 1.
 DR SMART; SM00332; PP2C; 1.
 DR SMART; SM00331; PP2C_SIG; 1.
 DR PROSITE; PS01022; PP2C; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 383 AA; 41899 MW; 9B50D7BCD5217B81 CRC64;
 Query Match Similarity 68.3%; Score 43; DB 10; Length 383;
 Best Local Similarity 71.4%; Pred. No. 45;
 Matches 5; Conservative 74; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 PRWMPCTL 9
 Db 14 PRWMPCTV 20

RESULT 13
 09FRJ1 PRELIMINARY; PRT; 458 AA.
 ID 09FRJ1; AC 09FRJ1;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 49.5 kDa protein.
 GN OSJNBB0064P21_12.
 OC Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzeae; Oryza.

THIS PAGE BLANK (USPTO)